Defense, Safety and Security
Kampweg 55
3769 DE Soesterberg
P.O. Box 23
3769 ZG Soesterberg
The Netherlands

www.tno.nl

T +31 88 866 15 00
F +31 34 635 39 77

TNO report

TNO 2017 R11325

Adaptive Maritime Automation: Final Report

Date	January 2018
Author(s)	Dr. J. van den Broek, Dr. J. van Diggelen, Dr. R. van der Kleij, Drs. T.F. Hueting, J. van der Waa, J.A. van Schendel MSc, J.J. Langefeld.
Number of pages	120 (incl. appendices)
Number of appendices	2
Sponsor	Early Research Program Human Enhancement
Project name	Adaptive Maritime Automation
Project number	060.27051

Managementuittreksel

Titel :	Adaptieve Maritieme Automatisering: Eindrapport
Auteur(s) :	Hans van den Broek, Jurriaan van Diggelen, Rick van der Kleij, Tom Hueting, Jasper van der Waa, Jef van Schendel, Anja Langefeld.
Datum :	januari 2018
Opdrachtnr. :	Early Research Program Human Enhancement
Rapportnr. :	TNO 2017 R11325

Door voortschrijdende technologische innovatie in de laatste decennia worden steeds meer taken van de mens overgenomen door geautomatiseerde systemen en robots. Dat geldt ook voor taken waarvoor het tot voor kort niet mogelijk leek dat ze volledig geautomatiseerd konden worden uitgevoerd, zoals het besturen van een auto of het varen van een schip (van den Broek, 2017).

Volgens Sarter, Woods en Billings (1994) is automatiseringstechnologie oorspronkelijk ontwikkeld om de precisie, prestaties en efficiëntie van werkprocessen te vergroten. Tegelijkertijd zou de werkdruk verminderen en zouden de opleidingseisen voor de operator kunnen worden aangepast. Ook werd het als technisch mogelijk beschouwd om autonome systemen te ontwikkelen die weinig, zo niet geen enkele menselijke betrokkenheid vergen en op die manier de kans op menselijke fouten te verminderen of elimineren.

Echter, hoog-geautomatiseerde systemen en zelfs volledig autonome systemen moeten beschouwd worden als een gezamenlijk mens-automatisering systeem. Aangezien 'autonomie' verwijst naar *zelfsturing* en *zelfredzaamheid*, is het mogelijk, zelfs wanneer de automatiseringsniveaus hoog zijn, dat ze tekortschieten wanneer de complexiteit van de omgeving toeneemt of wanneer de automatisering faalt (Van den Broek, Schraagen, Te Brake, & Van Diggelen, 2017). Deze gezamenlijke mens-automatisering interactie vormt ironisch genoeg een fundamenteel automatiseringsprobleem, namelijk:

> *Hoe meer het werk wordt geautomatiseerd en hoe betrouwbaarder en robuuster de automatisering is, hoe kleiner de kans dat de menselijke operator, die overzicht moet houden over de geautomatiseerde systemen, zich bewust is van kritische informatie en de controle handmatig kan overnemen wanneer dat nodig is (Endsley, 2016).*

Ondanks de hoge graad van automatisering speelt de persoon die het systeem monitort (operator, bestuurder, piloot) daarom een cruciale rol. Het is uiteindelijk deze bedienaar die de controle van het geautomatiseerde systeem moet kunnen overnemen wanneer de automatisering om welke reden dan ook faalt.

De concentratie (vigilantie) die het monitoren van hoog-geautomatiseerde autonome systemen vergt is in het algemeen voor mensen moeilijk op te brengen. Studies laten zien dat het vigilantieprobleem een vorm van zelfgenoegzaamheid (complacency) met zich meebrengt. Complacency houdt in dat een operator geleidelijk steeds minder geneigd is constant te willen weten wat de toestand van het systeem is en zich makkelijker overgeeft aan het gevoel dat alles wel goed gaat. Het kan ertoe leiden dat de operator het systeem te veel gaat vertrouwen (overreliance), ook in situaties waarin dat eigenlijk niet kan en die dus potentieel gevaarlijk zijn (Bradshaw et al., 2013). Dit effect staat bekend als het out of the loop-handelingsprobleem (Endsley & Kiris, 1995; Kaber & Endsley, 2004). Het heet ‘out of the loop’, omdat de operator in onvoldoende mate onderdeel is van het bedien- en beheersproces met als gevolg dat de situation awareness (SA) niet toereikend is om effectief te kunnen ingrijpen.

De benadering die in het project Adaptive Maritime Automation (AMA), onderdeel van het Early Research Program Human Enhancement (ERP-HE), wordt gevolgd is dat de functieallocatie niet dichotoom is, dat wil zeggen dat of de automatisering in control is of de operator, maar het mens-automatisering systeem zich aanpast met een meer geschikte samenwerkingsstijl naar gelang de situatie.

Het uitgangspunt van het gezamenlijke mens-automatisering paradigma, dat de manieren vaststelt waarop de operator en de automatisering samenwerken, is dat het een aanzienlijke invloed zal hebben op de SA van de operator en de prestaties bij het toezicht houden op de automatisering en ingrijpen indien nodig. Naast de fysieke displays (visueel, auditief of tactiel) die de operator ten dienste staan, zijn er verschillende fundamentele aspecten van het systeem die bepalen hoe interactie tussen mens en automatisering plaatsvindt, hoe rollen en verantwoordelijkheden ertussen worden verdeeld en hoe vaak deze toewijzingen zullen veranderen. Vaak besteden systeemontwikkelaars nauwelijks aandacht aan deze ontwerpbeslissingen; het automatisering interactie paradigma heeft echter een aanzienlijk effect op de complexiteit van het systeem en de mate van betrokkenheid en werklast van de operator, die allemaal het systeemtoezicht en de interventie aanzienlijk beïnvloeden. (Endsley, 2016).

Om meer kennis te ontwikkelen over het effect van ontwerpbeslissingen op de complexiteit van het systeem, de mate van betrokkenheid en werklast van de operator, en om de toegevoegde waarde van het gezamenlijke mens-automatisering paradigma te demonstreren hebben we een ondersteuningstool ontwikkeld en toegepast voor een maritieme use case, namelijk dynamic positioning systemen. Deze systemen worden toegepast voor operaties waarbij schepen langere tijd met grote precisie op een bepaalde positie gehouden moeten worden, waarbij operators 24/7 de systemen monitoren om in te grijpen in geval het schip van de aangegeven positie dreigt te verliezen. De dynamic positioning use case is gekozen omdat het huidige systeemontwerp alle bovenbeschreven kenmerken heeft waardoor de kans klein is dat de menselijke operator, die overzicht moet houden over het autonome DP-systeem, zich bewust is van kritische informatie en de controle handmatig kan overnemen wanneer dat nodig is.

Dynamic positioning systemen

Dynamic positioning systems (DP-systemen) zijn computergestuurde systemen die gebruikt worden om automatisch de positie van een schip vast te houden of ervoor te zorgen dat het schip een vooraf bepaald traject vaart. DP-systemen worden voornamelijk toegepast in de offshore-industrie onder andere bij het baggeren, het leggen van kabels en pijpen, het uitvoeren van duik- en booroperaties en het oppompen van olie. Schepen met DP-systemen worden steeds groter en complexer. Incidenten hebben daarom ook steeds grotere consequenties.

Floating production, storage, and offloading (FPSO) schepen bijvoorbeeld, zoals de Oaka Mizu, van Bluewater die is afgebeeld, worden lange tijd (maanden tot jaren) door middel van een DP-systeem recht boven een olieput gepositioneerd om via een flexibele zuigbuis olie op te pompen. De uitdaging is om een FPSO-schip ondanks wind, golven en stromingen met een marge van enkele meters op positie te houden. Als het schip te veel afdrijft, kan de zuigbuis breken, met alle gevolgen van dien (loss of position incident). Een loss of position incident (LOP-incident) veroorzaakt aanzienlijke kosten als gevolg van productieverlies en eventuele materiële en milieuschade. Het laatste redmiddel bij een incident is het gecontroleerd afkoppelen van de zuigbuis, waarmee breuk- en milieuschade worden voorkomen.

Het grootste deel van de tijd kan het DP-systeem de gewenste positie vasthouden, maar ondanks dat, monitoren vier DP-operators in ploegendienst 24/7 het systeem om eventueel verlies van positie te voorkomen.

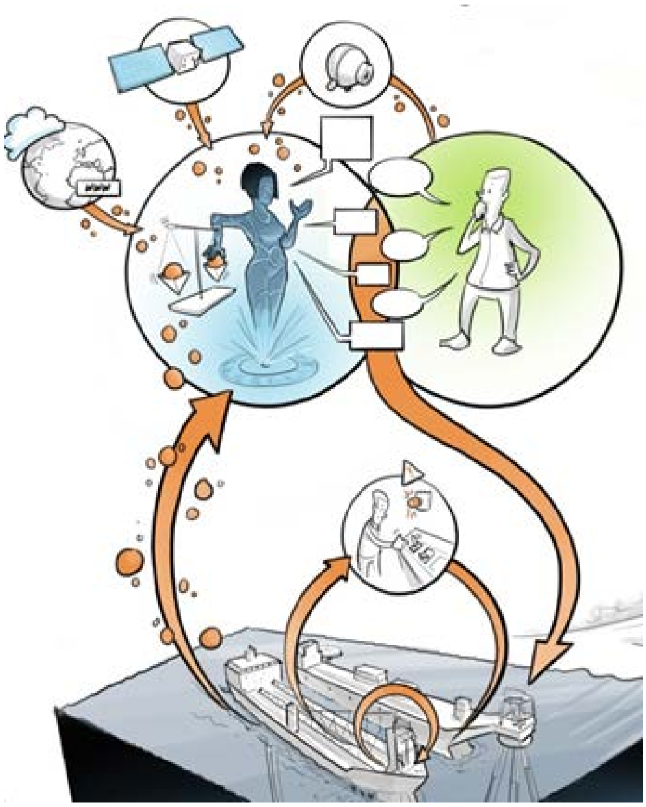
Het effect van de hoge graad van automatisering van het systeem is dat de DP-operator een superviserende rol heeft. Dat betekent dat operators langdurig geconcentreerd en oplettend moeten zijn zonder dat er veel gebeurt zodat de werklast (te) laag is. Het gevolg is dat de SA vaak onvoldoende is om snel en effectief in te kunnen grijpen, wat onnodig tot LOP-incidenten leidt (Van der Kleij, Te Brake, en Van den Broek, 2015).

Hierbij komt dat DP-systemen complex zijn en moeilijk te doorgronden. Anders dan in een auto bijvoorbeeld, kan de DP-operator een probleem vaak niet direct waarnemen maar slechts afleiden uit de informatie die via de interface (conning display) van het DP-systeem ontsloten wordt. Het zoeken naar informatie en instellingen kost tijd. Dit kan problematisch zijn, omdat de beschikbare tijd om te reageren op een drift off (LOP door wind of stroming) of een drive off (LOP door eigen aandrijving) in het algemeen zeer kort is en de kans op het voorkomen van een LOP snel afneemt nadat de fout optreedt. Het detecteren van het probleem, het identificeren van de fout, het bedenken van een oplossing en het implementeren van de oplossing kosten vaak te veel tijd om een LOP te voorkomen, zeker omdat het schip ook nog tijd nodig heeft om te reageren.

Intelligente ondersteuning van de DP-operator.

De doelstelling van AMA is om, vanuit het automatisering interactie paradigma, kennis over het effect van ontwerpbeslissingen op de complexiteit van het systeem, de mate van betrokkenheid en werklast van de operator te ontwikkelen.

Om interactie mogelijk te maken hebben we een intelligent operator support system (IOSS) ontwikkeld dat medieert tussen DP-systeem en operator, als een virtuele team genoot, c.q. agent.



Waar de eerste twee control loops (van binnen naar buiten) in de tekening representeren de huidige superviserende manier van werken:

1. de operator bewaakt de sensoren en metingen van het DP-systeem, en
2. de operator reageert op alarmen gegenereerd door het DP-systeem.

De derde interactieve control loop wordt gevormd door het IOSS dat functioneert als een intelligente virtuele teamgenoot die de operator op basis van verschillende informatiebronnen kan ondersteunen. Het IOSS maakt gebruik van zogenoemde slimme meldingen (smart notifications) om met de operator te communiceren. Slimme meldingen bestaan uit berichten in dialoogvorm, zoals ook wordt toegepast bij sms-berichten.

Het creëren van een interactief Intelligent Operator Support Systeem op basis van verschillende AI-technieken maakt nieuwe functionaliteiten mogelijk ten opzichte van de bestaande mens machine interface (MMI):

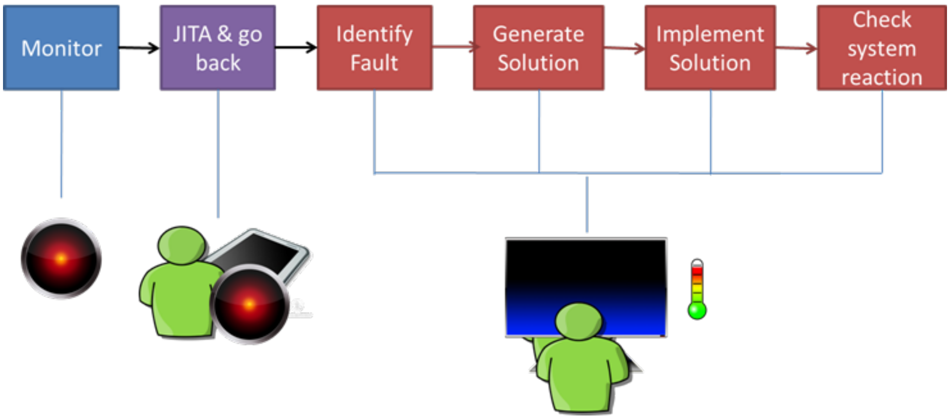
- Omgaan met onzekerheden: AI-algoritmen produceren vaak voorspellingen met een zekere mate van onzekerheid. Aangezien alarmen gebaseerd moeten zijn op solide feiten (en hieraan procedures worden gekoppeld), zijn ze niet geschikt om onzekere afleidingen, aannames en andere verhandelbare zaken te communiceren.
- Betekenisvolle communicatie: Alarmen met één regel tekst bevatten beperkte informatie zonder de mogelijkheid om grafische visualisaties te gebruiken, wat niet voldoende is om de redeneerlijn van een AI-algoritme te verklaren.
- Mobile toepassing: Besturingsinterfaces voor toezicht vereisen veel schermruimte en zijn geoptimaliseerd voor werkstations. Ze zijn niet geschikt voor mobiele apparaten.
- Data integratie: Gegevensvisualisatie is beperkt tot eigen scheepsgegevens. Andere informatiebronnen (bijvoorbeeld van internet of andere bronnen) worden niet weergegeven en worden niet in aanmerking genomen.
- Grotere monitoring capaciteit: Als de hoeveelheid (sensor) gegevens toeneemt, kan niet alles worden gemonitord omdat er hier onvoldoende ruimte voor is en dit ook niet voor mensen begrijpelijk is.
- Verbeterde visualisatie: Meterachtige gebruikersinterfaces zijn niet geschikt voor het communiceren van (onzekere) gegevensafgeleiden

Situation Awareness Recovery
Het vigilantieprobleem zoals hierboven beschreven was aanleiding om na te denken over mogelijkheden om de operator te ondersteunen bij het opbouwen en onderhouden van SA. De conclusie is:

a) het is niet mogelijk om de operator te vragen een constant niveau van SA te onderhouden in situaties die intern (het DP-systeem) en extern (omgevingscondities) stabiel zijn, en
b) het is niet zinvol, omdat niet objectief vastgesteld kan worden wat een adequaat niveau van SA is in een stabiele statische situatie.

De vraag of de SA van een operator adequaat is, kan alleen beantwoord worden als de context, c.q. de situatie waarbinnen gehandeld moet worden gegeven is. Pas dan kan bepaald worden welke specifieke informatie een operator nodig heeft om de situatie te begrijpen en te beslissen op welke manier het beste gehandeld kan worden. De taakuitvoering, maar ook het individuele probleemoplossend vermogen van de operator is bepalend voor de informatie die nodig is (en waarnaar dus gezocht wordt) om in control te zijn. Samenvattend betekent dit dat een operator moet kunnen 'schakelen' van een globaal SA-niveau (voldoende om een mogelijke ontwikkeling te kunnen zien aan komen, SA level 1) naar een specifiek SA niveau die dusdanig adequaat dat de operator de situatie en de consequenties ervan begrijpt (SA level 2) en dat het inzicht dusdanig tijdig wordt opgebouwd dat adequate handelend kan worden opgetreden (SA level 3). Het te vermogen om context specifieke en adequate SA op te bouwen vanuit een toestand van globale SA, noemen we SA recovery.

De figuur hieronder geeft schematisch weer welke rol het IOSS (het ronde element) speelt bij SA recovery.



Het monitoren van het DP-systeem en de omgevingscondities met als doel het detecteren van ‘state changes’ die mogelijk tot een LOP-situatie kunnen leiden, wordt overgenomen door het IOSS. Als de toestand van systeem en omgeving stabiel zijn (geen state changes), is het voor de DP-operator niet nodig actief zijn SA te onderhouden. Dit gegeven biedt de operator de mogelijkheid een secundaire taak te verrichten op de brug of ‘roaming’ te zijn en zich elders op het schip op te houden.

Als de DP-operator de brug verlaat, heeft deze in het roaming concept de beschikking over een portable tablet. Doordat op de tablet de belangrijkste informatie van dat moment grafisch wordt weergegeven, blijft de globale SA van de operator actueel. De DP-operator draagt tevens een smart watch. Als de situatie normaal is en stabiel, laat de smart watch van de operator een rustellende ‘hartslag’ zien doormiddel van het rustig pulseren van de gekleurde ringen. Als het IOSS reden ziet om de operator ergens van op de hoogte te brengen, gebeurt dat door de smart watch te laten trillen en wordt de hardslag frequentie verhoogd. Op de tablet wordt uitgelegd wat er aan de hand is. Doordat het IOSS de operator op de hoogte brengt van mogelijk belangrijke veranderingen, wordt in feite de SA recovery van de operator in gang gezet. De operator kan op basis van deze informatie begrip opbouwen van de situatie en bepalen of zich een serieus probleem aandient en er noodzaak is om naar de brug terug te keren of niet.

Eenmaal terug op de brug wordt de SA recovery verder ondersteunt door alle context specifieke informatie aan de operator aan te bieden alsmede de ‘state change’ informatie, dat wil zeggen aan te geven wat is er veranderd in de tussentijd dat de operator van de brug was. Wat dit roaming concept betekent voor de operationele praktijk van de operator hebben we doormiddel van een video gevisualiseerd (TNO, 2016).

Is het niet tegenstrijdig dat de íronie van automatisering’ wordt tegengegaan met de toevoeging van nog meer automatisering? Nee dat is het niet. Ten eerste, omdat we niet gebruik maken van traditionele automatisering. De toegevoegde interactieve control loop wordt namelijk gevormd door agent technologie waarmee de SA recovery van de operator wordt ondersteund op de momenten dat dat nodig is. Ten tweede zorgt de technologie die de monitoring mogelijk maakt ervoor, dat veel meer databronnen betrokken kunnen worden waardoor mogelijke verstoringen in een zeer vroeg stadium opgemerkt kunnen worden, en deze zogenaamde zwakke signalen aan de operator gecommuniceerd kunnen worden waardoor de anticipatietijd toeneemt. Op de derde plaats wordt een interactie concept toegepast dat gericht is op het overbrengen van informatie en het ontwikkelen van SA.

Intelligent operator support system.
De volgende functionaliteiten zijn ontwikkeld en geïntegreerd in het IOSS:

1. *human aware artificial intelligence*, om de operator te volgen in wat hij doet en niet doet en daarop de ondersteuning aan te passen;
2. *explainable artificial intelligence,* om mee te denken met de operator en uitleg te geven;
3. *data analytics*, om de omgeving te monitoren en het systeem te superviseren;
4. *procedurele ondersteuning*, om werkafspraken te maken en procedures te ondersteunen.

Human aware artificial intelligence
Het real time 'volgen' van het gedrag van de DP-operator is een vorm van human aware AI. Human aware betekent dat een systeem 'kennis' opbouwt van de gebruiker en op basis van die gegevens de ondersteuning aanpast.

De simpelste vorm van human aware AI is het detecteren van de locatie van de DP-operator. Die kan vastgesteld worden met behulp van locatiesensoren. De afstand die de DP-operator moet afleggen om naar de brug te komen en de aard van zijn bezigheden op dat moment, kunnen worden meegenomen in de timing van het moment dat de DP-operator van informatie wordt voorzien. Afhankelijk van de plek waar de operator is en de taak waar hij mee bezig is, kan het systeem de operator meer of minder tijd geven om weer op zijn plek te komen.

Explainable artificial intelligence
De tweede IOSS functionaliteit die het project voor ogen heeft, is dat de ondersteuning zich gedraagt als een virtuele teamgenoot die meehelpt, oplet, waarschuwt, meedenkt en uitleg geeft. Zo'n virtuele teamgenoot moet een herkenbaar voorkomen hebben, een avatar. Het beroemdste (of meest beruchte) voorbeeld van een virtuele teamgenoot is de Microsoft Office Assistant, waarvan de meest herkenbare representatie (avatar) Clippy werd genoemd. Het is met Clippy niet goed afgelopen, omdat deze niet goed genoeg was afgestemd op wat mensen in een bepaalde situatie aan ondersteuning nodig hebben. De fout die Clippy maakte was dat hij te vaak en op de verkeerde momenten de aandacht trok en daardoor hinderlijk werd. Ook viel Clippy gebruikers lastig met trivialiteiten terwijl deze met andere, meer serieuze taken bezig waren. Kortom, Clippy bezat niet de juiste eigenschappen die nodig zijn voor een goede samenwerking: het vermogen om een model van de gebruiker en de taakcontext te construeren.

Het vermogen van systemen om te kunnen uitleggen wat de situatie is en te kunnen aangeven wat een systeem wel en niet kan doen, wordt aangeduid als explainable AI. Explainable AI is technologie die de noodzakelijke transparantie, begrijpelijkheid en voorspelbaarheid van complexe geautomatiseerde systemen verbetert.

Het IOSS maakt gebruik van zogenoemde slimme meldingen (smart notifications). Slimme meldingen bestaan uit berichten in dialoogvorm, zoals ook wordt toegepast bij sms-berichten. De gebruiker kan irrelevante meldingen inactief maken en vragen om van relevante kennisgeving op de hoogte gehouden te worden. Het systeem kan zich op deze manier leren instellen op wat de operator wenselijk en niet wenselijk acht.

Data analyse

De bewakingsfunctie van het IOSS is niet alleen van belang als mogelijke oplossing van het vigilantieprobleem van de DP-operator. In vergelijking met mensen kunnen met computerkracht veel meer variabelen gevolgd en gecombineerd worden. Dat gebeurt met data-analysetechnieken (data analytics). Met data-analysetechnieken wordt in grote datasets (big data) gezocht naar statistische verbanden. Zo kan het IOSS de kans bepalen dat een probleem zich gaat voordoen als bepaalde kenmerken aanwezig zijn. Het systeem ‘leert’ de verbanden te leggen op basis van gebeurtenissen, omstandigheden en zwakke signalen die in het verleden aanwezig waren toen problemen ontstonden. Op deze manier is het mogelijk om in een vroeg stadium, nog voordat alarmbellen afgaan, problemen te zien aankomen of althans te weten dat de kans daartoe reëel aanwezig is. Hierdoor wordt de DP-operator zich al in een vroeg stadium bewust van een mogelijk probleem en heeft hij meer tijd om zich voor te bereiden en naar oplossingen te zoeken. In het kader van de FPSO-case, zijn data-analysetechnieken toegepast om de kans op een loss of position incident te voorspellen.

Het probleem van data-analysetechnieken is dat ze statistische verbanden leggen die voor mensen vervolgens lastig te interpreteren zijn. Welke actie onderneem je bijvoorbeeld als je weet dat er 10% kans op regen is: neem je dan wel of geen paraplu mee? Er moet dus nog veel onderzoek gedaan worden naar de manier waarop de resultaten van data-analysetechnieken geïnterpreteerd en uitgelegd moeten worden. Aan de andere kant, het feit dat wordt aangeven dat er een kans bestaat dat een bepaalde combinatie van factoren tot een probleem kan leiden, zal de operator alert maken en zal hij gestuurd door de verwachting actief naar informatie gaan zoeken. Als er informatie wordt gevonden die de ontwikkelingen bevestigen geeft hem dat de mogelijkheid voorbereidingen te treffen om voortijdig te anticiperen op de te verwachten situatie.

Er moet veel data beschikbaar zijn om voldoende betrouwbare voorspellingen te krijgen. Als bijvoorbeeld een verband wordt gevonden op basis van een enkele meting, dan is het voorspellend vermogen gering; het kan dan namelijk gewoon toeval zijn. Echter, als het verband op basis van duizenden metingen wordt vastgesteld, is de voorspelling veel overtuigender en betrouwbaarder. Ook de betrouwbaarheid van de voorspelling zal moeten worden geadresseerd, want als het systeem te vaak onterecht alarm slaat, zal het vertrouwen in het systeem afnemen. Wellicht reageren operators dan vervolgens minder alert op een volgende waarschuwing, terwijl er in dat geval wel echt iets aan de hand kan zijn.

Als antwoord hierop is een intuïtieve zekerheidsmaat ontwikkeld (Intuitive Certainty Measure, ICM). ICM berekent de kans dat gegeven output correct is.

Procedurele ondersteuning

Een aspect van samenwerking is dat leden van een team ‘werkafspraken’ met elkaar maken over wie onder welke omstandigheden welke taken wel of niet mag uitvoeren. Analoog hieraan is het mogelijk om werkafspraken tussen de operator en het IOSS te maken. De werkafspraken kunnen bijvoorbeeld betrekking hebben op acties die het systeem zonder toestemming mag uitvoeren (en daarna gemeld worden), acties waar het systeem vooraf toestemming voor nodig heeft en taken die primair bij de operator blijven liggen. Ook is het mogelijk afspraken te maken over hoe vaak en wanneer de DP-operator geïnformeerd wil worden (zie afbeelding 7). Deze functionaliteit van werkafspraken kan gebruikt worden om onder bepaalde omstandigheden de werkdruk van de operator op een aanvaardbaar niveau te houden en niet gestoord te worden door trivialiteiten.

IOSS gebruikers evaluatie

Om de functionaliteit en bruikbaarheid van de IOSS voor de FPSO-praktijk te verifiëren, werd deze geëvalueerd door verschillende ervaren DP-operators (DPO's). In samenwerking met een DP-instructeur van het Scheepvaart en Transport College (STC) Rotterdam, werden hiervoor twee FPSO-specifieke scenario's gecreëerd waarin ooit een storingsmodus werd geïntroduceerd.

Vijf ervaren DPO's werden uitgenodigd om deel te nemen aan het onderzoek om het IOSS op een kritische manier te evalueren door middel van een semi-gestructureerd interview. De operators werden gerekruteerd via professionele netwerken, o.a. die van de onderzoek partners Bluewater B.V. en de STC-groep. De operators die aan de evaluatie hebben meegewerkt hebben operationele ervaring in verschillende DP-domeinen, waaronder maar niet beperkt tot steenstort, kabelleggen, boren, offshore constructie en militaire operaties voor de Koninklijke Marine.

De operators die deelnamen aan de evaluatie werden uitgenodigd op onze onderzoeksfaciliteit in Soesterberg. De evaluatie gebeurde met één operator tegelijk en duurde ongeveer 4 uur per persoon. De opzet bestond uit verschillende systemen die aan elkaar gekoppeld waren. Eén monitor werd voor de operator geplaatst, gekoppeld aan een computer met het Dynamic Positioning-systeem, ontwikkeld door RH marine (Conning 4500). Deze computer diende ook als een serverplatform voor het IOSS-systeem dat was verbonden met het RH-systeem voor maritieme DP. Op dezelfde computer werden ook de simulaties uitgevoerd en deze informatie werd ingevoerd in het DP-systeem en de IOSS. Een tabletcomputer (zie foto) werd gebruikt om de IOSS (web interface) weer te geven. De operators kunnen deze tablet in de roaming conditie meenemen en door de beschikbare informatie bladeren.

Om de functionaliteit van de IOSS aan te testen, werden twee scenario’s ontwikkeld. Beide scenario's begonnen met een stationaire DP-operatie bij rustig weer. Na een periode van 10 minuten actief monitoren werd de DPO door het IOSS geïnformeerd dat het veilig was om de DP-desk gedurende ten minste de volgende 15 minuten te verlaten en de supervisie over het DP-systeem aan het IOSS over te laten. Op dit punt werden de DPO's van het DP-systeem vandaan genomen en hielden ze via een tablet-computer toegang tot het IOSS. De DPO’s mochten de IOSS gebruiken om SA te behouden, maar werden ze niet geïnstrueerd om dit te doen. In het scenario was voorzien dat er na enige tijd nadat de operator niet actief met het systeem bezig was (roaming conditie) een fout zou gaan optreden. In het eerste scenario was dit een plotselinge een plotselinge toename van de windsnelheid. In het tweede scenario was dit een thruster-storing (dat wil zeggen een probleem met de aandrijving van het schip).

Op het moment dat de storing optrad, of verwacht werd te gaan optreden in het geval van een toename van de windsnelheid, zou de IOSS de operator informeren en moest deze reageren op de situatie. Dat beide scenario's begonnen met een stationaire operatie tijdens kalme zeeën was noodzakelijk om de operators de gelegenheid te geven te gaan ‘roamen’ ten behoeve van onze evaluatie.

De evaluatie was gericht op verschillende elementen, zoals de waarop de IOSS zichzelf voorstelt aan de operator, de meldingen van de IOSS, bruikbaarheid van de verschillende ondersteunende functies, vertrouwen in het systeem en tot slot de algemene indruk van het IOSS.

De reacties van de operators op het IOSS en het feit dat ze van hun DP-station konden weglopen (roaming) waren positief. Meerdere operators zagen "de toegevoegde waarde" en dat het "absoluut" potentie had. Een van de operators zei dat "Dit zou mijn werk een stuk leuker hebben gemaakt". Ze erkenden dat het interessant zou zijn voor bedrijven om "te besparen op mensen, om geld te besparen", maar belangrijker nog, ze zagen mogelijkheden voor DP-operators om andere taken uit te voeren in plaats van het DP-systeem te bewaken. Monitoring van DP was tijd "die op een nuttiger manier kon worden besteed" en het zou de taak van de DP-operator "aangenamer" maken.

Meerdere operators voegden echter voorwaarden toe aan deze positieve verwachtingen zoals het belang van betrouwbaarheid en voegden eraan toe dat een dergelijk systeem nuttig zou zijn "in een ideale wereld" en dat het "echt waterdicht moet zijn". Ze merkten op dat een dergelijke verandering in nauwe samenwerking met de industrie zou moeten worden doorgevoerd en zou ook gevolgen hebben voor de opleiding van operators.

Een belangrijk element van de IOSS is de intuïtieve zekerheidsmaat (ICM). De Operators na elke kennisgeving met inbegrip van een ICM gevraagd hoe zij het hebben ervaren. In eerste instantie gaven de operators aan geen behoefte te hebben om te worden geïnformeerd over de zekerheid van het advies van het IOSS. Naarmate het scenario vorderde en ze werden geconfronteerd met het feit dat het IOSS fouten kon maken en geloofden de operators dat een bepaalde zekerheidsaanduiding erg belangrijk zou zijn om het vertrouwen in het systeem te behouden na een onzekere voorspelling. Kort gezegd, het IOSS mocht fouten maken, op voorwaarde dat het de operator vooraf waarschuwt over de waarschijnlijkheid dat er een fout optreedt.

Het concept van het IOSS werd door de operators overwegend goed ontvangen. Er was een algemene overtuiging dat een systeem als het IOSS van groot voordeel zou kunnen zijn tijdens DP-achtige operaties. Er zijn enkele opmerkingen gemaakt die wijzen op het belang van vertrouwen tussen mens en machine, zodat een succesvolle samenwerking mogelijk wordt. Hoewel de operatoren van mening waren dat het IOSS goede informatie verstrekte, werd de behoefte gevoeld om de suggesties van het IOSS dubbel te controleren en te vergelijken met hoe de operator de situatie zelf zou beoordelen. Dit geeft het vertrouwen van de operator in het systeem weer. Het is belangrijk om te zorgen voor een passend niveau van vertrouwen om verkeert gebruik van automatisering te voorkomen. Uiteindelijk moet het doel zijn dat de mens weet in welke situaties het systeem betrouwbaar is. Dit kan worden bereikt door de ervaring van de operator, maar ook door assistentie van de ICM.

Adaptive Support for Supervisory Control Operators in Highly Reliable Automated Systems.

Een tweede vorm van human aware AI is het meten en ondersteunen van de situation awareness van de DP-operator. Op welke manier dat vastgesteld kan worden is erg afhankelijk van de context, gedacht kan worden aan eye tracking. Een eye-trackingsysteem bestaat uit kleine camera's die de kijkrichting van de ogen registreren. Op deze manier kan bijvoorbeeld nauwkeurig vastgesteld worden hoe lang en naar welke elementen van een conning display (interface) de DP-operator kijkt. Op basis van deze eye-trackingdata kan vervolgens geanalyseerd worden welke gegevens de DP-operator kennelijk belangrijk vindt in een bepaalde situatie. Als de DP-operator bijvoorbeeld allerlei verschillende datatypes in ogenschouw neemt, zou dat een indicatie kunnen zijn dat hij 'zoekende' is en nog begrip van de situatie aan het opbouwen is. Als de DP-operator echter vooral naar een bepaald data type kijkt, zegt dat hij begrip heeft van de situatie. Zou het niet mooi zijn als het 'technische systeem' op die momenten zijn situation awareness, als ondersteuning, kan delen met de DP-operator?

Semi-autonome systeemprestaties zijn sterk afhankelijk van hoe operator en automatisering als een team functioneren omdat beide componenten sterk van elkaar afhankelijk zijn. Tussen handmatig beheer en volledige automatisering kunnen verschillende niveaus van automatisering of samenwerkingsvormen worden onderscheiden. Een speciale vorm van samenwerking tussen mens en automatisering zijn adaptieve systemen. Adaptieve systemen zijn systemen waarbij de locus of control in de loop van de tijd varieert. Dit houdt in dat de verantwoordelijkheid voor een specifieke sub taak verschuift van de automatisering naar de operator of vice versa. Adaptieve automatisering is een subset van adaptieve systemen, wat impliceert dat de automatisering taken van de menselijke operator overneemt wanneer de behoefte zich daarom voordoet.

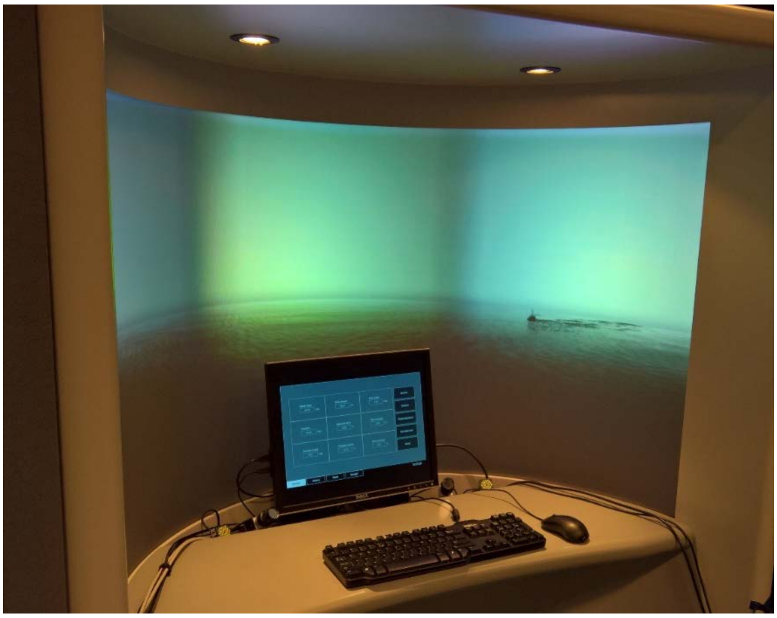
In het onderzoek naar *Adaptive Support for Supervisory Control Operators in Highly Reliable Automated Systems* was de vraag of adaptieve automatisering kan worden gebruikt om problemen van cognitieve onderbelasting en potentiële out-of-the-loop (OOTL) prestatieproblemen, met name verlies van SA, op te lossen. Dit is nog niet eerder onderzocht. Meer specifiek onderzoeken we de potentie van een adaptief algoritme dat beoordeelt of de operator oplet (of niet) en beslist, op basis van de noodzaak van betrokkenheid van de operator, of de operator moet worden gewaarschuwd, c.q., getriggerd. Het gebruik van een dergelijke hybride triggeringstrategie met behulp van prestatiemeting om te bepalen of een operator al dan niet oplet, werd ook niet eerder bestudeerd.

Het onderzoek is opgezet om te onderzoeken of adaptieve automatisering bijdraagt aan een hogere systeemefficiëntie in vergelijking met situaties zonder adaptieve automatisering (d.w.z. geen adaptief algoritme). We zijn ook geïnteresseerd om te zien of deze adaptieve ondersteuning extra voordelen zou kunnen bieden in vergelijking met statische ondersteuning, bijvoorbeeld in prestaties of verminderde werklast.

Op basis hiervan zijn drie hypotheses geformuleerd:

1. We voorspellen dat prestatie- en situatiebewustzijn wordt verbeterd met adaptieve ondersteuning, terwijl de algehele mentale belasting zou worden verminderd en dat deze voordelen specifiek zouden worden geassocieerd met adaptieve (in tegenstelling tot statische) ondersteuning.
2. We voorspellen dat stress een negatief effect heeft op de prestaties en het bewustzijn van de situatie.
3. We voorspellen dat de effecten van adaptieve ondersteuning meer uitgesproken zijn bij hoge niveaus van stress in vergelijking met lage niveaus van stress.

Voor de experimenten is een DP-simulatie ontwikkeld, ontworpen om een deel van de cognitieve vereisten te isoleren die verbonden met DP-systemen. Het ontwerp van de simulatie was gebaseerd op echte DP-systemen en feitelijke sensorinformatie. Scenario's waren gebaseerd op een analyse van kritieke gebeurtenissen. Wanneer zich een kritieke gebeurtenis voordeed, moesten de deelnemers beslissen welke actie het best kon worden ondernomen. Toen een beslissing werd genomen, stopte het scenario. De taak werd tijdens trainingssessies getraind in criterium en alle vereiste informatie voor een beslissing werd tijdens het experiment beschikbaar gesteld om de werkgeheugenbelasting te verlichten. Op de foto is de experimentele opstelling afgebeeld.



De proefpersoenen moesten vier taken uitvoeren, die werden gedefinieerd als primaire en secundaire taken. De primaire taak was om de veilige werking van het geautomatiseerde DP-systeem te bewaken en te reageren op gevaarlijke situaties. De secundaire taken waren: (a) het informeren van de management van waarden van sensoren op regelmatige tijdsintervallen, het meten van voorspellende geheugenprestaties; (b) een beheerderstaak (in eigen tempo) die specifieke vooraf bepaalde woorden markeert in een tekstbestand en (c) om bedreigingen (binnenkomende schepen) op zee te detecteren en te rapporteren, zichtbaar op het gesimuleerde buitenaanzicht. Eén schip zou tegelijkertijd verschijnen. Alle schepen begonnen hun nadering vanaf dezelfde afstand die onzichtbaar was voor de bestuurder. Ze naderden met gelijke snelheden, maar vanuit verschillende hoeken, tussen -70 en +70 graden.

We hebben de hypotheses getest in een gemengd twee factoren ontwerp. Ondersteuning werd gemanipuleerd voor elke proefpersoon en had vier niveaus: geen ondersteuning, statische ondersteuning, adaptieve ondersteuning op basis van kritieke gebeurtenissen; en adaptieve ondersteuning die wordt opgeroepen door een hybride triggeringstrategie. Stress werd gemanipuleerd tussen proefpersonen en had twee niveaus: hoge stress en lage stress.

In totaal namen zesentwintig studenten als óperator' deel aan deze studie. Eén deelnemer werd uitgesloten vanwege een fout van de computer om de gegevens op te slaan. De uitkomsten van de studie zijn gebaseerd op de resterende vijfentwintig deelnemers (13 mannen, 12 vrouwen). Hun leeftijd varieerde van 19 tot 44 jaar (M = 27,5, SD = 7,8). Elke deelnemer had normaal gehoor en gezichtsvermogen. Geen van hen had eerdere ervaring met dynamische positionering, maar sommigen hadden deelgenomen aan een voorstudie. Deelnemers kregen € 45 voor deelname en er werd nog eens € 40 beloofd voor de best presterende en € 20 voor de tweede best presterende deelnemer om de motivatie te verbeteren.

Het eerste doel was om een adaptief algoritme te ontwerpen dat beoordeelt of de operator oplet (of niet) en beslist, op basis van de situatie, of de operator moet worden gewaarschuwd (of niet). Uit onze gegevens bleek dat we op basis van oogbewegingen van de operators succesvol waren in het vaststellen of de operator aandacht schonk aan de primaire taak en dus ondersteuning bood wanneer dit nodig werd geacht.

We waren echter minder succesvol in het aantonen van de voordelen van adaptieve automatisering voor de algehele systeemprestaties. Vanwege een onvoorzien probleem met logboekreacties konden we niet testen op de effecten van ondersteuning op de responstijd van de operator, onze primaire prestatiemaatstaf. Andere metingen toonden geen verschillen tussen de niet-ondersteunde conditie en de hybride adaptieve ondersteuningstoestand. Er werden geen effecten gevonden van ondersteuning op mentale inspanning of situatiebewustzijn. Onze studie bevestigde echter de nadelige effecten van slecht ontworpen ondersteuning op de prestaties: de statische en de kritieke omstandigheden voor eventondersteuning, die beide een overvloed aan alarmen veroorzaakten, de meeste valse, verminderden de algehele systeemprestaties. De conclusie zou daarom kunnen zijn dat hybride adaptieve ondersteuning de operators heeft belet om overbelast te raken met alarmen, waardoor hij of zij stabiele systeemprestaties kan handhaven op de relatief korte duur van de taak.

Een interessante bevinding is de afwezigheid van stress-effecten op de prestaties. Dit lijkt op het eerste gezicht verrassend. Het kan er echter op wijzen dat operatoren zich aanpasten aan de stressor door effectieve compenserende strategieën te gebruiken. Volgens de theorie van Hockey (1997) worden compenserende strategieën doorgaans gevonden om door inspanning verhoging de aandacht te verleggen van secundaire naar primaire taken of het nemen van risico's te vergroten. Echter, post-hoc analyses lieten geen bewijs zien dat deelnemers met succes coping mechanismen aanpasten tijdens scenario's met hoge stress. Er werden geen verschillen gevonden tussen stressomstandigheden in de tijd die de deelnemers aan de primaire taak hadden toegewezen in vergelijking met de tijd die aan secundaire taken was besteed. Een mogelijke verklaring is de complexiteit van de taak. Sommige deelnemers klaagden dat de taak te complex en moeilijk was om goed te presteren. We merkten inderdaad op dat deelnemers op een laag niveau presteerden en misschien geen ruimte lieten voor verdere prestatieverminderingen als gevolg van stress.

Operator Cognitive State modelling using unobtrusive measures

Om op het juiste moment adaptieve geautomatiseerde ondersteuning te kunnen bieden, is het belangrijk dat de automatisering zich bewust is van de cognitieve toestand van de menselijke operator. Context specifieke adaptieve automatisering is belangrijk om zogenaamde 'disruptieve automatisering' te voorkomen, het fenomeen waarbij ondersteuning alleen maar bijdraagt aan de toch al hoge werkbelasting en het vermogen van de operator om correct te reageren verder belemmert. Het vermogen om de operator state te kunnen modelleren wordt human aware computing, c.q. human aware AI genoemd: de automatisering bouwt een beeld op van het doen en laten van de operator en stemt daar de aan te bieden ondersteuning op af.

Het Operator Cognitive State modelling experiment is bedoeld om vooruitgang te boeken in het modelleren van operator-state. De experimenten zijn uitgevoerd met DP-trainees in de DP-trainingssimulator van het Scheepvaart en Transport College (STC) in Rotterdam, tijdens een reguliere training. Het experiment was dusdanig opgezet dat de trainees tijdens de training geen last ondervonden (non unobtrusive). Het experiment stopte voor de dag dat examens werden afgenomen (laatste dag van de trainingsweek) om elke vorm van beïnvloeding te voorkomen.

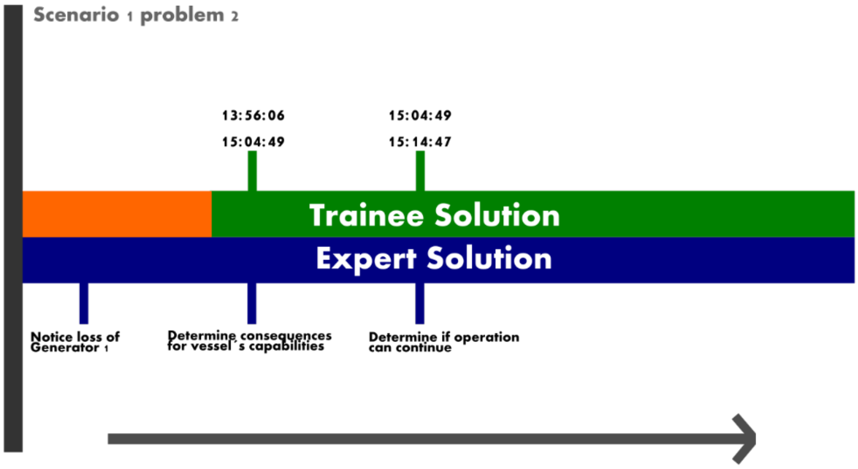
Tijdens de trainingssessies werden de DP-operator en zijn brugploeg geconfronteerd met verschillende problemen (failure modes). Om te bepalen of de Situation Awareness van een operator adequaat voor het probleem, is het nodig om het handelen van de operator te vergelijken met de optimale oplossing zoals uitgevoerd door een domein expert. Om dat te kunnen realiseren was het nodig drie uitdagingen op te lossen:

1. het vastleggen van het probleemoplossend gedrag van de operator;
2. het formeel beschrijven van de expertoplossing;
3. het vergelijken van probleemoplossend gedrag van de operator met expertoplossing.

Om het probleemoplossend gedrag van de operator te kunnen vastleggen was het belangrijk om alle systeeminteracties en de communicatie die relevant zijn voor de DP-taak te registreren en te loggen. Dit betrof de directe interacties van de DPO met het DP-systeem, zoals de schermen die door de operator werden geselecteerd, de specifieke informatie waar de operator naar keek, maar ook de DP-instellingen die werden gewijzigd. Tweede bron van informatie betrof interactie met zijn brugbemanning tijdens het bespreken van relevante informatie en mogelijke handelwijzen. Tenslotte was er ook communicatie tussen de brug en externe partijen (bijvoorbeeld machinekamer, boorplatform) die de DPO van vitale informatie zouden kunnen voorzien. Al deze informatiebronnen werden gelijktijdig vastgelegd en gesynchroniseerd ten behoeve van de analyse. Op de foto is de installatie van het eye tracking systeem zichtbaar (de camera's en infrarood sensors op het railsysteem)

De optimale oplossingspaden, de tweede uitdaging, werden vóór de trainingssessies ontwikkeld in samenwerking met de DP-instructeur die ook de training leidde. Alle storingsmodi die tijdens de training zouden worden gepresenteerd, waren bij de experimentele leider van tevoren bekend. Er is een Cognitieve taakanalyse gemaakt voor al deze faalwijzen. Dat wil zeggen dat er op een cognitief niveau is gekeken naar informatie waarop bepaalde handelingen zouden moeten volgen (if-then relaties) of andersom welke informatie rechtvaardigt het wel of niet uitvoeren van een actie. In totaal zijn er negen expert reacties beschreven op evenzoveel probleemsituaties.

Uiteindelijk is van alle geïntroduceerde probleemsituaties een analyse gemaakt van de traineeoplossing t.o.v. de expertoplossing. Beide oplossingen zijn in een grafische beschrijving samengebracht. Het plaatje hieronder geeft daar een voorbeeld van. De blauwe balk geeft de expert oplossing weer met de informatie elementen die een actie hadden moeten triggeren. Het oranje deel geeft aan dat de trainee het kritieke informatie element wel gezien heeft maar later dan gemoeten, dan wel gekund had. Het groene deel geeft aan dat de acties erna conform de expertoplossing zijn. Rood wordt gebruikt om aan te geven dat of informatie gemist is of dat geen of een verkeerde actie is uitgevoerd.



Het experiment heeft aangetoond dat het mogelijk is de cognitieve toestand van de DP-operators te meten zonder dat ze daarvan hinder ondervinden bij de taakuitvoering. Door verschillende informatiebronnen te integreren was het mogelijk om te bepalen wat de operators op een bepaald moment aan het doen waren. Hierdoor kon het onderscheid worden gemaakt tussen de vraag of een operator het probleem eenvoudigweg niet opmerkt (niet zo opvallend) of dat hij actief bezig was met oplossingsstrategieën. Dankzij de eye-tracker was het ook mogelijk om te observeren welke informatie de operator zocht tijdens het oplossen van het probleem. Hierdoor kon een oplospad worden gereconstrueerd en vergeleken met dat van een expertoperator. Het oplos pad gaf inzicht in wanneer en hoe een reeks fouten begon. Ontwikkelaars van toekomstige ondersteuningssystemen kunnen deze inzichten gebruiken om te bepalen wanneer en hoe ondersteuning geboden moet worden, d.w.z. tijdens detectie of beoordeling van het probleem.

Uit de prestaties van de operators werd duidelijk dat verschillende oorzaken van fouten konden worden onderscheiden. Hoewel de precieze aard van de fouten varieert, zijn er overeenkomsten tussen deze fouten. Zo hadden bijvoorbeeld meerdere fouten kunnen worden voorkomen als het DP-systeem meer opvallende alarmen had gegeven. De operators hebben bijvoorbeeld een ‘generatorfout’ gemist omdat ze de daling in stroomtoevoer niet hebben opgemerkt, ook al was deze zichtbaar op hun scherm. De zichtbaarheid van het alarm kan worden verbeterd door het ontwerp van de DP-interface.

Hoewel sommige fouten kunnen worden voorkomen door eenvoudigweg het ontwerp van de DP-interface te verbeteren, waren de meeste fouten niet het gevolg van mislukkingen die onopgemerkt bleven, maar het resultaat van het feit dat de operators niet in staat waren de juiste maatregelen te treffen. Dit soort fouten bieden toekomstige systemen een mogelijkheid om operators te ondersteunen wanneer hun taken zeer veeleisend zijn. De human aware technieken die in dit rapport worden beschreven kunnen dienen als basis voor het verfijnen van de modellering van de operator state, te beginnen met het vervangen van post-analyse door (bijna) real time analyse van de operator state. Op deze manier kan een toekomstig ondersteuningssysteem adaptieve ondersteuning ook op gebied van het oplossen van problemen.

Eindevent

Op 16 november 2017, is er in het Spant in Bussum een zogenaamd eindevent georganiseerd om de kennisproducten die het ERP programma heeft opgeleverd aan marktpartijen en toekomstige gebruikers te demonstreren. Ter gelegenheid van dit eindevent is een brochure ontwikkeld (TNO, 2017). Hieronder een foto van de demo's die het project Adaptive Maritime Automation het kader van dit eindevent heeft verzorgd.

Tot slot willen wij RH Marine, Bluewater en het Scheepvaart en Transport College (STC) Rotterdam hartelijk danken voor het beschikbaar stellen van software en domeinexpertise maar nog meer voor de prettige en vruchtbare samenwerking.

Contents

Appendices

A Raw interview data

B Optimal solution path for each failure mode

1 Introduction

This is the final report of a three year research project in adaptive maritime automation, which started in January 2015 and ended December 2017. The project was part of the Early Research Program Human Enhancement that initially consisted of three projects, namely adaptive automotive automation (AAA), human resilience, and adaptive maritime automation (AMA). In 2017 a project on cyber security was added.

From the start, the focus of AMA was on the human factors issues relating to the monitoring and controlling of highly automated systems, in particular dynamic positioning (DP) systems. DP systems are computer-controlled systems that automatically keep a floating vessel at a specific position or to follow a pre-defined path (tracking) by using its own propellers and thrusters. Applications include shuttle tanker operations, deep water drilling (drilling rigs), diving and ROV support operations, dredging and rock dumping, pipe laying and pipe trenching operations, cable lay and repair operations, but also military operations (e.g., mine countermeasures) (see also Fossen, 1994). The number of vessels with DP systems has increased in recent years. This is due mainly to increased oil and gas exploration at sea, as well as offshore operations, such as wind turbine construction, drilling, diving support, and anchor handling.

From a human factors point of view DP systems are a valuable object of research because DP systems are basically highly automated autonomous systems, taking over tasks previously performed by people, with the intention of increasing safety, accuracy, and reliability (see also Parasuraman, Mouloua, and Molloy, 1996; Sheridan, 1992; Wickens, 1998). When automation is introduced into a system, or when there is an increase in the autonomy of automated systems, developers often assume that adding automation is a simple substitution of machine activity for human activity (Woods and Sarter, 2000). Empirical data on the relationship of people and technology suggest that this is not the case and that traditional automation has several negative performance and safety consequences associated with it stemming from the human out-of-the-loop (OOL) performance problem (Endsley and Kiris, 1995; Kaber and Endsley, 2004). The essence of highly automated autonomous systems is, that the operator has no direct need to constantly know what the status of all parts of the system is, because the system is controlling all components itself. Only after a failure arises the operator needs to take over this task and take appropriate action(s) to prevent the failure from harming the operation, or abort the operation in time to prevent accidents. Consequently, the low SA due to a high level of automation makes that the operator cannot intervene quickly and effectively if the automation fails. This is known as the OOL-performance problem, as the operator is not an active part of the process, (Parasuraman, Molloy and Singh, 1993; Tjallema et al., 2007). This is especially problematic in DP operations where the available time-window for reacting on a drive-off incident is in general very short, and the chances of preventing an accident decrease rapidly after the fault-initiation (Chen and Moan, 2003; Sandhåland et al., 2015).

Highly automated systems and even fully autonomous systems should be considered within a joint human-automation framework. Since ‘autonomy’ refers to self-directedness and self-sufficiency, it is possible, even when the levels are high, that they fall short when complexity of the environment increases or when automation fails. However, a joint human-automation framework possess significant challenge. An automation conundrum exists in which as the more automation is added to a system, and the more reliable and robust that automation is, the less likely that human operators overseeing the automation will be aware of critical information and able to take over manual control when needed (Endsley, 2016).

To overcome the automation conundrum the Adaptive Maritime Automation (AMA) project, part of Early Research Program Human Enhancement (ERP-HE), followed an approach in which the function allocation between human and automation is not dichotomous, i.e. that the automation is in control or the operator, but the human-automation system adapts with a more suitable style of cooperation for the situation. Hence, within a joint human-automation framework it is not about the operator taking over tasks from the automation, but adapting towards a more suitable collaboration style for the situation (Van den Broek, Schraagen, Te Brake, & Van Diggelen, 2017).

In addition to the physical displays (visual, auditory, or tactile) provided to the operator, there are several fundamental aspects of the system that determine how the human and automation will interact, how roles and responsibilities will be allocated between them, and how often these allocations will change, which Endsley (2016) terms *the automation paradigm*. Often system developers give scant attention to these design decisions; however, the automation interaction paradigm has a significant effect on the complexity of the system and the level of engagement and workload of the operator, all of which significantly influence system oversight and intervention.

Our ambition is to develop more knowledge about the impact of design decisions on the complexity of the system, the degree of involvement and workload of the operator, we have applied and tested the automation interaction paradigm on the basis of a maritime use case: the application of dynamic positioning systems for operations in which ships have to be held at a certain position with great precision for a longer period of time. Hence, dealing with human – automation interaction issues was not the principle goal of the project but a means to develop state of the art knowledge of joint human-automation frameworks.

In order to demonstrate the benefits of a joint human-automation framework for DP-practise, we developed, together with the industry, a transparent (human-in-the-loop) adaptive automation platform that substantially improves safety for manoeuvring and control tasks capable of assessing the operator’s need for support, based on the system, environment, and the current and predicted operator’s functional state, that is, the variable capacity of the operator for effective task performance in response to task and environmental demands.

1.1 The stationary DP use case

In order to a to develop technology for a transparent (human-in-the-loop) adaptive automation platform we chose an FPSO (Floating Production Storage and Offloading) use case. For this purpose we worked together with Bluewater Energy services BV that operates a fleet of five FPSO vessels. FPSO installations are oil tankers that mine and store crude oil. The oil is regularly loaded into a shuttle tanker for transport. FPSOs can be brought quickly to new operations, so it is very useful for small oilfields and to operate the first wells before a final platform is ready. Critical is the positioning above a well. Figure 1-1, depicts the MUNIN, one of the Bluewater FPSO vessels.

Figure 1-1. MUNIN, a Bluewater FPSO vessels.

FPSO vessels are kept in position (stationary) for a long time (months to years) right above the oil well to collect oil via a flexible suction tube (see figure 1-2). An FPSO vessel must therefore remain in the same position despite wind, waves and currents. This type of DP application is called stationary DP. If the ship drifts too much, the suction pipe may break. A so called loss of position incident (LOP incident) causes considerable costs as a result of production loss and possible material and environmental damage and needs te be avoided. The last resort in case of an LOP incident is the controlled disconnection of the suction pipe, which prevents breakage and environmental damage. Most of the time, the DP system can hold the desired position. Nonetheless, four DP operators work in shifts to monitor the system 24/7 to prevent any loss of position.

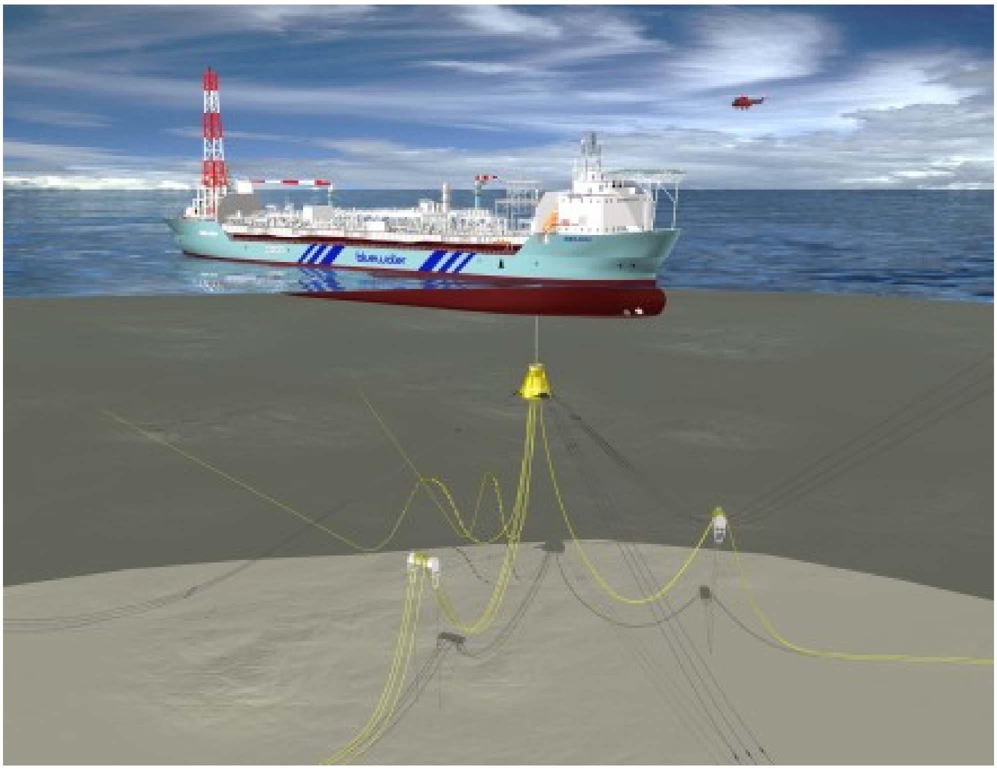


Figure 1-2. FPSO vessels are kept in position (stationary) for a long time (months to years) right above the oil well to collect oil via a flexible suction tube.

Despite the fact that DP systems are highly autonomous systems, the DP operator still plays a crucial role. Ultimately, it is the DP operator who has to take over control from the DP system to prevent LOP incidents in case the DP system fails. The role of the DP operator can be characterized as a supervising role. When supervising the work- and information load are low and it requires a constant high level of concentration (vigilance). A combination typically not suited for people. Hence, the concentration (vigilance) that monitoring a DP system requires is also difficult to uphold for a DP operator, especially at night. Studies by Parasuraman and Riley (1997) show that the vigilance problem entails a form of self-complacency. Complacency means that an operator is gradually becoming less and less inclined to constantly want to know what the state of the system is and to surrender more easily to the feeling that everything is going well. It can cause the operator to trust the system too much (overreliance), even in situations where this is not possible and which are therefore potentially dangerous.

The effect of the high degree of automation of the system is that the DP operator is often unable to take control of the system quickly and effectively (Van der Kleij, Te Brake, and Van den Broek, 2015). This effect is known as the out-of-the-loop problem of performance (Endsley and Kiris, 1995; Kaber and Endsley, 2004). It is called 'out of the loop' because the operator is insufficiently part of the operating and control process.

DP systems are complex systems and are difficult to grasp. The DP operator can only derive problems or critical situations from the information accessed via the interface (conning display) of the DP system. Searching for information and settings takes time. This can be problematic, because the available time to respond to a drift off (LOP by wind or current) or a drive-off (LOP by own propulsion) is generally very short and the chance of a LOP occurring quickly decreases after the error occurs (Chen and Moan, 2004 Sandhåland, Oltedal, Hystad and Eid, 2015). Tjallema, Van der Nat, Grimmelius and Stapersma (2007) endorse this by stating that detecting the problem, identifying the error, coming up with a solution and implementing the solution often take too much time to implement, especially since the ship also needs time to respond.

1.2 Roaming operator concept

The vigilance problem was a reason to investigate the possibilities for taking over the monitoring task by 'technology'. The idea is that this gives the DP operator the possibility to leave his workstation on the bridge to undertake other activities like administrative work or rest. The fact that the DP operator can leave the bridge and undertake other activities makes the job more interesting. Because the technology provides the DP operator the opportunity to 'roam', the approach has been given the name 'the roaming concept'. The roaming concept is explained and illustrated through a concept video (TNO, 2016).

Figure 1-3. The DP operator carries a smart watch. The information is displayed on a tablet device.

The system developed within the project is called the intelligent operator support system (IOSS). In this system, the DP operator has access to a tablet that he takes with him when he leaves the bridge. The critical information of that moment is displayed graphically on the tablet. The DP operator also carries a smart watch (figure 1-3). If the situation is normal and stable, the operator's smart watch shows a quiet "heartbeat" by pulsing the colored rings. If the IOSS sees reason to inform the operator about something, this is done by vibrating the smart watch and the tablet will explain what is going on.

The implementation of the roaming concept requires that different techniques be developed and brought together in the IOSS. The starting point in the development of the IOSS is that the DP operator is supported but that he himself continues to do the real thinking; his role is therefore not taken over by 'technology' but is supported by it. In fact the IOSS.

To make the roaming concept possible, the following functionalities have been developed and integrated into the IOSS:

1 Human aware artificial intelligence, to follow the operator in what it does and does not do and to adjust the support accordingly;
2 Explainable artificial intelligence, to think along with the operator and explain;
3 Data analytics, to monitor the environment and to supervise the system;
4 procedural support, to make working arrangements and support procedures.

In chapter 2, the Artificial Intelligence (AI) technologies which enable the roaming operator concept are described and explained.

To verify the functionality and usefulness of the IOSS for the FPSO practice, it was evaluated with the help of several experienced DPO's. For this purpose a human in the loop demonstrator has been developed (see Figure 1-4 for an impression).

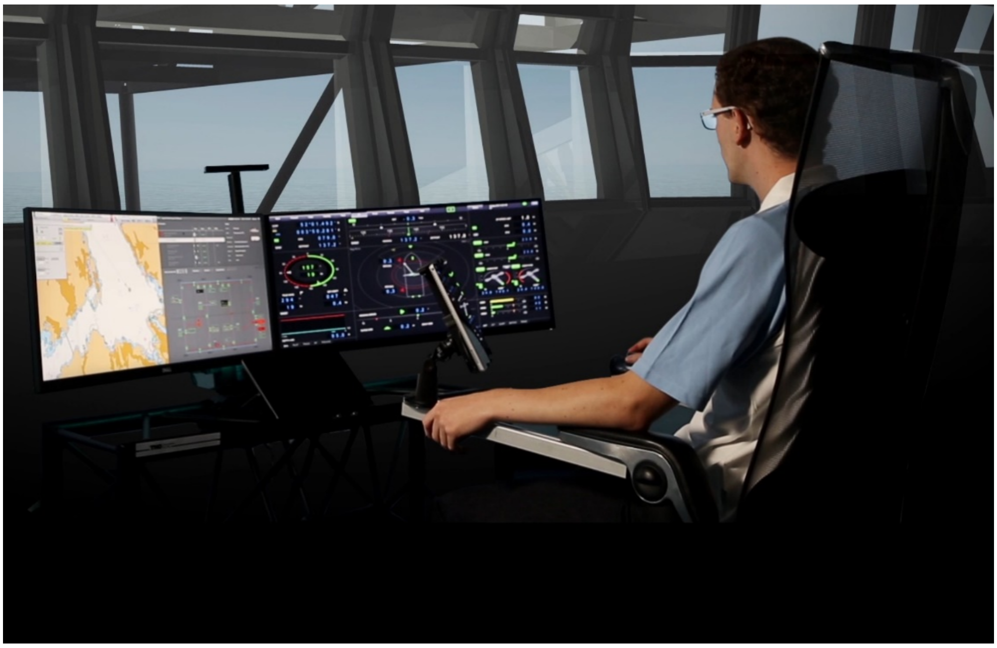


Figure 1-4. The IOSS on the bridge.

Chapter 3, provides a detailed overview of the IOSS user evaluation which was conducted at the research and simulation facility of TNO at Soesterberg.

1.3 Adaptive support

An important operator variable for safe and reliable DP operations is situation awareness (Heinen, 2016), or SA in short. It is important that the operator's level of SA is maintained at high levels. The FPSO performance is highly dependent on how operator and automation function as a team. Both components are highly interdependent. Between manual control and full automation, different levels of automaton, or collaboration forms, can be distinguished. In order to investigate whether adaptive automation helps in achieving higher overall system efficiency, as compared to situations without adaptive automation (i.e. no adaptive algorithm) a research on Adaptive Support for Supervisory Control Operators in Highly Reliable Automated Systems, was set up. The setup and results of the research effort are explained in chapter 4.

1.4 Human Aware AI

An important aspect of Intelligent Operator Support is the real time measuring and supporting the situation awareness of the DP operator. This is called human aware AI or human aware computing. A simple instantiation of human aware computing is detecting whether an operator is sitting behind the desk or is roaming. A more complicated form of human aware AI is to determine what the SA of an operator is and if it fits the task context.

In the DP training simulator of the Shipping and Transport College (STC) in Rotterdam, in 2016 and 2017 experiments in human aware AI were carried out by two Rotterdam Miniport Institute students (Poelman, 2016; Houtkoop, 2017). In short, the experiment consisted of digitally recording the interaction of DP trainees with the DP interface during several DP training sessions. The question was whether the analysis of the data could provide insight in how the trainees deal with a DP problem (a so-called failure mode) that the instructors introduce at a given moment in the simulation. The question to be answered consists of two parts: 1) it is possible to deduce from the interaction data how DP trainees deal with the failure mode and 2) can the approach taken by the DP trainees be compared with the ‘optimal’ problem approach provided by the instructors (expert solution)? If the comparison shows that what the trainees do deviate from the expert solution, this may be an indication of an inadequate situation awareness.

If this can be established, the IOSS could be augmented with a decision support functionality, by giving (extra) instructions or pushing context relevant information, for example. Such functionality is very similar to the interaction between instructors and DP trainees in existing training situations. To put the DP trainees on the right track, instructors say something like: 'Have you taken this dot that into account ...?'

The experimental design and outcomes are described in chapter 5.

1.5 Acknowledgement

The research on adaptive maritime automation has been made possible through the sponsorship of the Early Research Program on Human Enhancement. We thank Bluewater, and especially Clemens van der Nat, for providing the project the necessary operational knowledge and subject matter expertise. We thank RHmarine for providing the DP conning system for experimental use and we owe Ehab el Amam a lot of gratitude for his technical help and discussion on the operation of DP-systems. We also thank the STC-group for use of the DP-training facility at Rotterdam for experimental use. We also thank Arie Goedknegt for help with establishing DP-scenarios and for providing us with valuable domain knowledge.

2 AI support technologies

To develop the AI support technologies which enable the roaming operator concept, we applied the situated Cognitive Engineering (sCE) approach that integrates technological, human factors (HF), and operational perspectives (Neerincx and Lindenberg, 2008). The four steps in the process are depicted in figure 2-1. As a result, the IOSS prototype consists of a number of combined software modules.

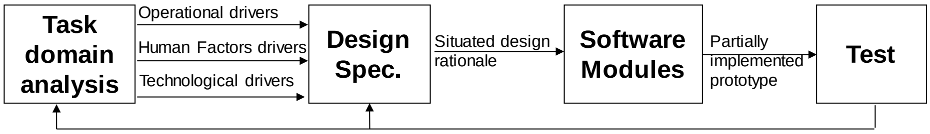


Figure 2-1. Steps in the situated Cognitive Engineering methodology.

In the first phase, we conducted a task domain analysis and identified the most important operational, human, and technological drivers. From a technological perspective, we have identified predictive analytics as a crucial technology to enable a roaming operator. Predictive analytics can be used to predict future situations based on data from the past using machine learning algorithms (Lent, Fisher, Mancuso, 2004). For example, to predict that environmental conditions are expected to remain stable, a system state allowing the operator to leave the bridge, or to predict when alarms are likely to appear, a system state requiring the operator to return to his workstation located on the bridge. From a human factors perspective, we identified a number of potential problems related to trust, cognitive overload, and other issues well known in the HF literature (Endsley, 2016). For example, misalignments of operator's trust in the system could occur because the performance of predictive analytics changes over time as more training data is used. These concerns have been adequately addressed in the design specification.

In the second phase, we elaborate the design specification aimed at providing a solution to the task domain problems. The results are recognizable as user requirements, design patterns, and claims (which specify the rationale behind a design decision).

The design specification is implemented in software modules in the third phase. We implemented the most important patterns and user requirements in an early phase of IOSS prototyping, enabling to test to whether or not they bring about the required results. Testing and evaluation of IOSS is the topic of the next chapter.

This chapter is organized as follows. We first discuss existing interface systems in section 2.1. Then, we discuss future technology in section 2.2. The architecture of IOSS will be discussed in section 2.3. In section 2.4, we zoom in on one particular aspect of this architecture, namely explainable AI for real time predictive analytics. Finally, the prototype of IOSS will be discussed in section 2.5.

2.1 Existing systems

Figure 2-2 shows the traditional DP interface in which the interaction with the human consists of two main activities:

- the human monitors the sensors and measurements from the DP system, and
- the human responds to alarms generated by the DP system.

Figure 2-2: Traditional DP interface.

This is a typical example of a supervisory control way of working. When this paradigm is applied to AI systems, we identify a number of shortcomings of this approach:

- AI algorithms often produce predictions with a certain degree of uncertainty. Since alarms should be based on solid facts (and have procedures attached to them that must be followed in case an alarm occurs), they are not suited to communicate uncertain derivations, assumptions, and other negotiable matters.
- Alarms containing one line of text contain limited information without the possibility to use graphical visualizations, which is not enough to explain the line of reasoning produced by an AI algorithm.
- Supervisory control interfaces require lots of screen space, and are optimized for work stations. They are not suitable for mobile devices (and therefore not for the roaming operator).
- Data visualization is limited to own ship data. Other information sources (e.g. from the internet, or other sources) are not displayed and not taken into account.
- If the amount of (sensor) data increases (which can be expected in a smart room environment), not everything can be monitored as there won't be sufficient space for this, nor will this be comprehensible for humans.
- Meter-like user interfaces are not suitable for communicating (uncertain) data-derivations

2.2 Towards a next generation operator support systems

2.2.1 Technological Drivers

In this section we identify a number of technology trends that will play a major role in the maritime domain and hence for future DP operations. Three of these technologies are outlined below.

Firstly, *predictive analytics techniques* are expected to have a major impact on the maritime world. One possible application is predictive maintenance where large quantities of sensor data are collected and used as input for a machine learning algorithm. Over time, the algorithm should be able to recognize system failures before they occur using historical data. Such a classifier would be useful for our DP application where potential component failures are important to the operator. Many other applications of predictive analytics to DP are conceivable, for example: predicting position-loss based on weather data or predicting operator's drowsiness based on physiological data (Singh, Bhatia, and Kaur, 2011 provide an example in the automotive domain).

Secondly, *Internet of Things* can be regarded as having a major impact in the maritime domain by allowing an unprecedented amount of data to be gathered and shared on a vessel (Lee, 2013). Virtually every component of a ship could become an information processing node in a large network. Applications in the DP domain could be monitoring the location of an operator, and disclosing vast amounts of additional information sources to the DP system to enable it to function more accurately.

Thirdly, computers are becoming more and more used as *personal assistants* (e.g., Siri[1], and google home[2]), which changes the relation between human and computer from that of a reactive tool to a more proactive entity, e.g. a teammate (Bradshaw et al, 2009). Also the IOSS should be viewed as a personal assistant and teammate.

Figure 2-3, shows the role of IOSS which is added as an extra layer of intelligent support on top of the existing DP-control system and alarm system.

[1] http://www.apple.com/ios/siri/
[2] https://madeby.google.com/home/

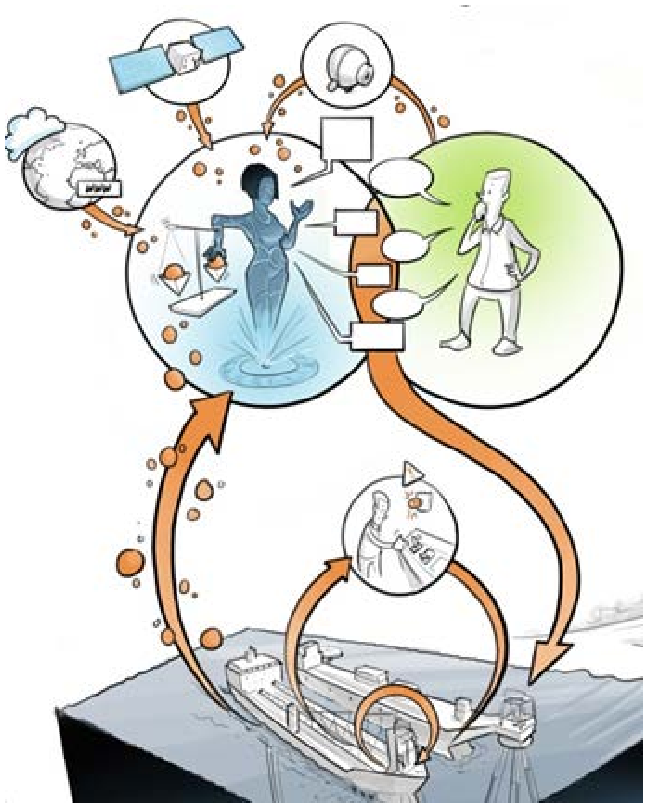


Figure 2-3. Control loops in the DP system. From inside to outside: (1) the DP control system, (2) the alarm system, (3) the agent support layer.

2.2.2 Design specification

The user requirements presented in table 2-1, are divided in five parts, each of which will be briefly discussed below.

Table 2-1. User requirements of IOSS.

Adaptive Automation
- IOSS should be adaptable w.r.t. task division and communication style
- IOSS should adapt its communication style according to user state
- IOSS should prevent cognitive overload of its user
- IOSS should behave according to a mixed initiative interaction style

User interface
- IOSS should support mobile and stationary UI's

Situational Awareness
- IOSS should support prediction of future situations
- IOSS should support change detection
- IOSS should support procedure awareness

Trust calibration
- IOSS should be able to explain itself
- IOSS should have a recognizable appearance

Agent architecture
- IOSS should be capable of acting in an open system
- IOSS should be capable of integrating information from multiple sources

The requirements for *adaptive automation* aim to ensure a balanced workload which is tailored to the current situation of the user. This impacts the density of information that is communicated between user and IOSS, and finding a proper balance is regarded to be a responsibility of both, i.e. mixed initiative interaction. This means that the user is capable of instructing the computer when and how it wishes to be notified about which information by making *working agreements* (Arciszewski, de Greef, and van Delft, 2009). The system also adapts its communication style to match the user's state, e.g. being brief when the operator is busy, and being more elaborate when the operator is not that busy.

The *user interface* requirements state that both mobile and stationary user interfaces are needed to allow the concept of a roaming operator.

The *Situation Awareness* requirements are intended to provide the operator with a sufficient level of SA. At their most fine grained level (not shown in table 2-1), these requirements specify exactly which information must be communicated in which types of situations. However, as stated above, these are adaptable to the user's preferences using working agreements.

The requirements regarding *trust calibration* are aimed to prevent distrust by ensuring that the agent is capable of explaining the outcomes of the predictive analytics algorithms, i.e. explainable AI (Lent, Fisher, and Mancuso, 2004). Because IOSS is used complementary to the DP-system (and its alarm system), a different trust relation should be built up with the DP system, and IOSS which learns over time and could mistakenly produce wrong predictions. To make it clear to the operator if he is interacting with IOSS or with the DP system, the IOSS must have a recognizable appearance. The last set of requirements deals with *architectural issues*, such as openness of the system, and access to digital information sources.

2.3 Technological architecture and design components

In order to translate the requirements stated above into technology, we use a modular approach where different requirements are implemented in separate modules. The overview of modules that make up the social layer between humans and DP system is depicted in the information flow diagram below:

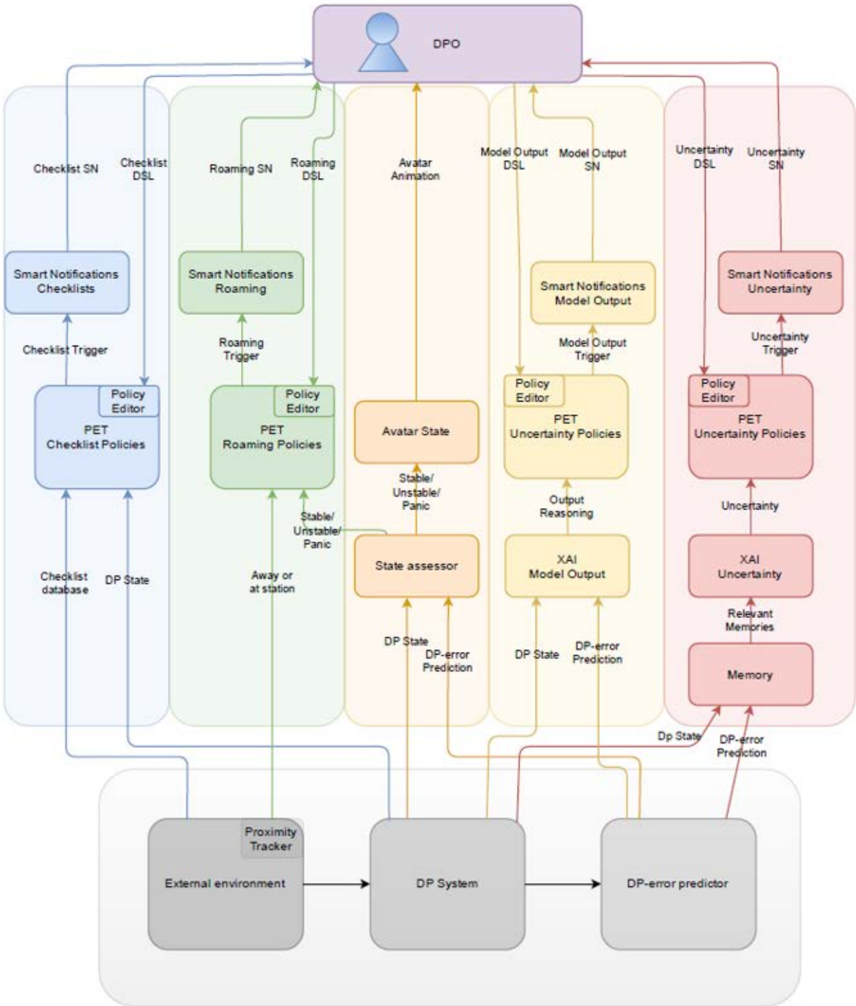


Figure 2-4. Overview of modules composing IOSS.

The essential aspects of these modules are smart notifications (which will be discussed in section 2.3.1), policy engine (which will be discussed in section 2.3.2), and explainable AI (which will be discussed in section 2.4).

2.3.1 Dialogues and smart notifications

Smart notifications are a set of design pattern for establishing interaction between humans and highly autonomous systems. The first design pattern aims to provide a solution for the problem that human control has some context requirements which must be fulfilled before control can be passed to the human. One of these context requirements is spatial location. For example, when a problem occurs while the system operates in fully autonomous mode, the human operator should be able to override from a workstation within a certain time limit.

Title	Demand operator to stay in vicinity of workstation.
Design Problem	The agent predicts that human intervention might be necessary soon, which cannot be done from a mobile device. It asks the operator to stay in the vicinity of the workstation.
Design solution	Popup window with short explanation of the type of expected problems and time frame. The operator can ask the agent for more explanation, and decide to agree or disagree to stay in the vicinity.
Use when	Agent expects to switch from autonomous mode to a semi-autonomous mode which requires a stationary operator.
Design rationale	Operator is more likely to follow the system's advice to stay in the vicinity if (s)he understands why this is necessary.
Example	
Status	Proto

The interaction design patterns such as the ones described above have been designed to realize sensible user interaction by themselves. This does not guarantee that the user can cope with multiple interactions running simultaneously. The next design pattern that we will discuss aims to solve that problem.

Title	Manage multiple interactions between user and system
Design Problem	When many interactions with the agent are required simultaneously, the user gets overloaded with information.
Design solution	A container window which contains all separate interactions as separate tiles. The important interactions are shown intrusively (i.e. in colour and large), and the less important interactions are shown non-intrusive (smaller and greyed out). The container shows the most important 7 windows in an intrusive way. The user can choose to dismiss any interaction as non-important using the "resolve" button.
Use when	Multiple different types of interactions are required simultaneously.
Design rationale	By limiting the amount of intrusive interactions to seven, human operators are capable of processing these simultaneously.

Title	Manage multiple interactions between user and system
Example	
Status	Proto

Smart notifications can be used to establish a dialogue between human and computer. These dialogues are structured using dialogue trees, such as the one shown in figure 2-5.

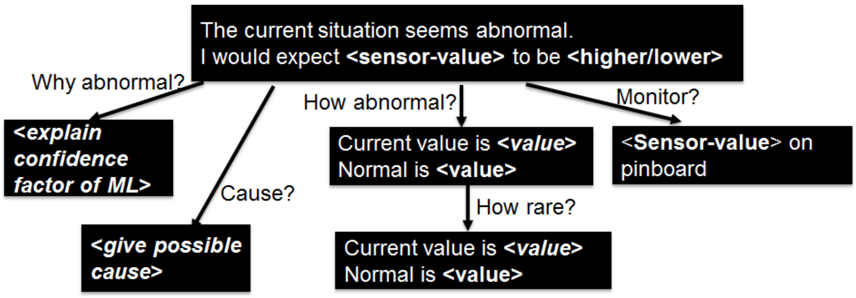


Figure 2-5. Dialogue tree.

The dialogue tree in this example can be used to provide various kinds of explanation on a weather predication.

2.3.2 Policy Engine TNO

As argued by Bunch et al. (2005), a policy engine can be used to specify notification rules that allow users to adapt when and how the user is notified. We follow a similar approach and specify policies in the Drools expert system language[3]. An example of a policy in our case is:

```
If wind speed is greater than 6 bft and the operator is
roaming then IOSS must suggest to the operator to come back
to stationary position
```

An important feature of our policy engine is that these rules are understandable for non-programming experts, which allows them to adapt these rules to their liking.

[3]https://www.drools.org/

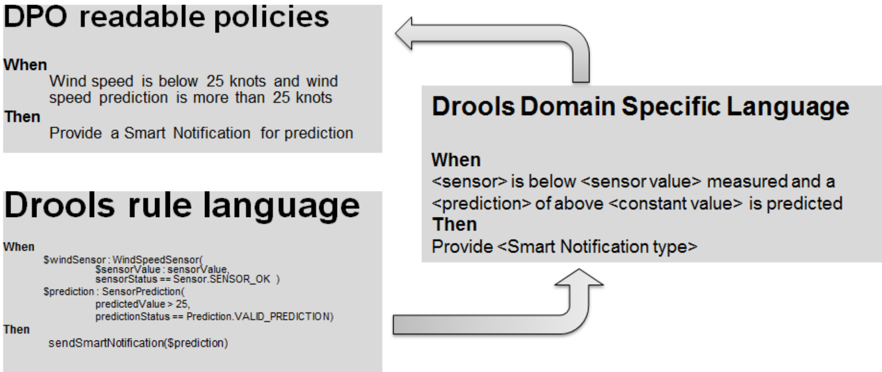


Figure 2-6. Implementation of a policy engine using drools.

The implementation of the Policy Engine TNO (PET) uses the Drools rule language, which can be directly translated to human-readable policies using a Domain Specific Language (DSL). An example is provided in the figure 2-6.

2.4 Explainable AI

The need for transparency and explanations towards end users of intelligent systems is becoming a necessity (Miller, 2017). This self-explaining capacity in intelligent systems allow them to become more effective tools that can maintain appropriate levels of trust. The research field of Explainable Artificial Intelligence (XAI) aims to develop and validate methods for this self-explaining capacity (Gunning, 2017).

The process of explaining requires at least two actors: the explainer (someone who explains) and the explainee (someone who receives an explanation).
Recent developments in XAI focus mainly on the system as the explainer and how it can generate explanations. It does not take the explainee into account, which is the 'receiving-end' of the system. A well designed explanation functionality needs to incorporate the user's wishes, context and requirements.

We developed a beneficial machine learning model for DP within IOSS, by adding an intuitive certainty measure (ICM) to aid the user in calibrating its trust in this model. Finally, we validated ICM in an experiment. In the next section, we describe the machine learning model and the way in which the required data was generated, the model was selected and trained. Finally, the intuitive certainty measure is explained, including the results of the qualitative experiment.

2.4.1 Realtime predictive analytics

The aim was to develop a machine learning model that can predict at least 15 minutes into the future whether the ship will drift or not, and if so, how large this drift will be. Because of lack of data to train such a model, data was generated using real weather data which was transformed to a fine-grain dataset using weather models with added sensor noise using sensor models. This resulted in a weather dataset with higher granularity; one data point per minute.

The weather dataset was used to generate scenarios of three hours that included a variety of weather situations; from clear weather to fierce storms and any transition in between.

These scenarios were fed to a ship simulation of which we extracted the simulated sensor data to retrieve a high-fidelity dataset of a ship's response to two years of weather data. This sensor dataset was used as the train and test set for the machine learning model.

The machine learning model selected was a Deep Neural Network. A neural net with three hidden layers (1536, 256 and 64 neurons respectively) was used with ReLu activation functions (Nair and Hinton 2010), trained using the ADAM optimizer (Kingma and Ba, 2015). No attempt was made to fully optimize the model when it reached an accuracy of 96.59% with hand tuned hyper-parameters on predicting three classes 15 minutes in the future; a drift of < 5m, between 5 and 10m or > 10m.

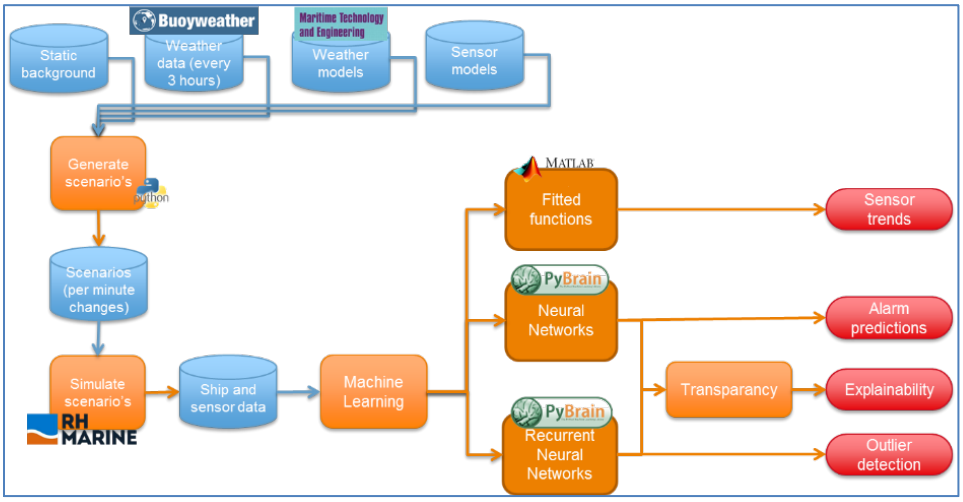


Figure 2-7. An overview of the machine learning process.

Figure 2-7, shows the process we applied to create the machine learning model to predict 15 minutes into the future whether the ship will drift from its stationary position. We started with a static background (e.g. GPS location), a weather dataset from a buoy, weather models from the literature and white noise models for sensors. These resulted in a large set of scenarios of 3 hours with a data point at every 1 minute. These scenarios were simulated in a ship simulation. From this simulation we extract the simulated sensor outputs on which we trained a machine learning model (a Deep Neural Network). On which we built the Intuitive Certainty Measure.

2.4.2 Intuitive Uncertainty measure

ICM computes the probability the currently given output is correct. It does this by weighing the difference of that output with the ground truths of a set of known past data points with the similarity of the current data point with those past data points. We visualized this in figure 2-8 for a simple example. Figure 2-8.a shows that the current data point is very similar to a number of other data points that have the same ground truths as the current output. Hence, ICM has a high certainty; it is likely that the current output will also be correct. Figure 2-8.b shows a different situation with a low certainty where it is more likely that the output is incorrect as the model made errors in the past for similar data points.

Finally, figure 2-8.c shows a situation where the data point is relatively new, such that none of the past data points has much impact on the certainty value resulting in a unknown certainty value.

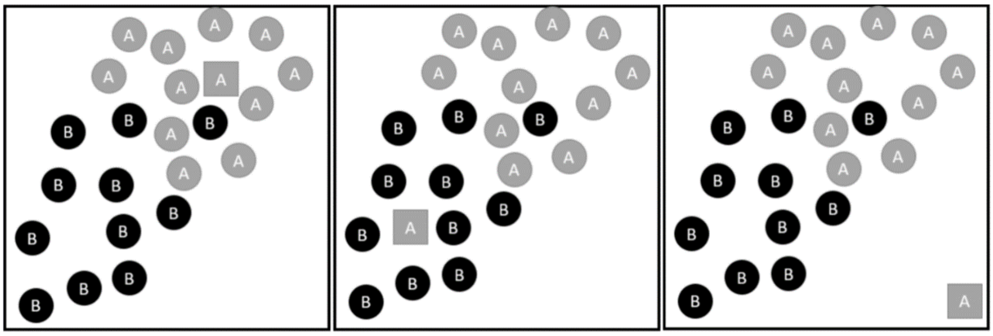


Figure 2-8. Three examples of how ICM works in a 2D binary classification task (class A or B) given a current data point with its output (square) and a set of known data points (circles) with their known ground truths or errors.

The intuitive explanation of this is that if the model often makes a mistake with a certain kind of input, ICM's certainty value will be low. If however it often provides a correct output, the certainty value will be high. Finally, if a data point is relatively new or on a decision boundary with equal densities on either sides, the certainty will reflect a 50/50 percent chance of being correct.

Note, that the choice of the similarity measure can be anything and is not restricted to Euclidean distance as this may not be appropriate given the data and model. The similarity measure, for example, may not be able to describe the relations between data points sufficiently for ICM to be accurate. However, Euclidean distance is relatively easy to understand, especially for numerical data with sensor information as the features (as in our use case). The eventual choice of the similarity measure is a trade-off between the desired performance of the measure and how well it can be understood by a non-expert.

ICM is based on the following three equations, with x as an arbitrary data point, M an arbitrary data set, d the used similarity function, s as the standard deviation used for the exponential weighting and M(T = A(x)) to select all data points in M with the same ground truth as the output of model A for x;

$$C(x \mid \sigma, M) = \frac{1}{Z(x \mid \sigma, M)} P(x \mid \sigma, M) \tag{1}$$

where P is the positive contribution to the certainty value:

$$P(x \mid \sigma, M) = \sum_{x_i \in M(T=A(x))} \exp\left(-\frac{d(x \mid x_i)^2}{2\sigma^2}\right) \tag{2}$$

and Z is the normalization constant:

$$Z(x \mid \sigma, M) = \sum_{x_i \in M} \exp\left(-\frac{d(x \mid x_i)^2}{2\sigma^2}\right) \tag{3}$$

The memory or dataset M is sequentially sampled according to three aspects. This strategy prefers data points with 1) a ground truth least common in the memory, 2) are sometime apart to mitigate any possible temporal dependencies and 3) are relatively dissimilar to all current points inside the memory. We refer to the original paper of ICM for a detailed description (van der Waa, van Diggelen, Neerincx and Raaijmakers, 2018). This memory is restricted to a fixed size, k, to prevent extreme computational costs. The number of computations increases exponential with each added data point and to store all data would quickly become unfeasibly for real world cases where the model and ICM may run for indefinite time.

The only parameter of ICM, besides the used similarity function d and memory size k, is s that is used for the exponential weighting. With this value the designer of ICM can determine a soft-threshold of the number of similar data points should be taken into account when computing the certainty value. The parameter is fairly robust, due to C being relative to all similarities but can be tweaked to handle sparse or dense data more appropriate.

ICM has several properties in common with other lazy learning techniques such as k-Nearest Neighbors (k-NN). In specific ICM is very similar to the weighted k-NN algorithm with an exponential weighting scheme where the normalization constant Z guarantees that all weights sum to one. ICM becomes an instance of weighted k-NN for non-linear regression with the model’s error as the dependent variable, the memory M to mitigate computation cost and an arbitrary distance function d.

2.4.2.1 Validation of the Intuitive Certainty Measure

In a small experiment we compared the understanding of three instance of different types of certainty measures by end-users; 1) ICM as a lazy learned meta-models, 2) the approach by Park et al. as active learned meta-models and 3) the soft-max output as a numerical output of the actual model. The experiment was done with a virtual smart assistant that supports an operator through situation predictions in a monitoring task. We simulated the operator’s work environment and the virtual smart assistant and provided realistic scenario’s and responses from the assistant including a certainty value for any made predictions. This simulation was used to get the participants acquainted with the assistant and the certainty values. It allowed us to ask beforehand if they felt it was useful to have a certainty value for each prediction.

This simulated work environment was followed by an interview where participants received increasingly more information about the three certainty measures.
The goal of the interview was to test if the participant understood a measure according to the ‘Comprehension’ level of Bloom’s taxonomy of learning (Forehand 2010). The interview went through several stages;

1. First stage
    a. Brief textual explanations of each measure and opportunity for the participant to rate the following;
        i. his understanding of the measure,
        ii. if he would define certainty like this,
        iii. whether the smart assistant should use this measure,
        iv. and any possible disadvantages he foresees with this measure (to test their ability to apply their understanding of the measure).

   b. Per measurement a moment for the participant to ask questions to allow the supervisor to rate for himself how well the participant seems to understand the examples.
   c. An explanation by the participant for each measure in their own words to rate by a machine learning (ML) expert on validity after the experiment.
2. Second stage
   a. Three concrete examples, both visual and textual, for each measure to illustrate its mechanisms where the participant could rate his level of understanding for each set of examples.
   b. For each set of examples a moment to ask questions to the supervisor, to allow the supervisor to rate for himself how well the participant seems to understand the examples.
   c. An explanation by the participant for each example in their own words to rate by a ML-expert on validity after the experiment.
   d. The participant's final preference for one of the three certainty measures and an explanation of what the smart assistant means with its provided certainty values, such that a ML-expert can validate whether the approach overlaps with one of the three measures.

The results of the five participants are shown in table 2-2, all were experts and potential end-users in the dynamic positioning use case. The two users that saw no use for a certainty measure believed that predictions should always be correct or otherwise not presented at all. All participants believed that they had some basic to advanced comprehension of each measure and its set of examples, however the experiment supervisor and ML-expert disagreed with this for both the 'numerical' and 'active learning meta-model' measures.

Table 2-2 Ratings of test subjects

	Participant:	P1		P2		P3		P4		P5		Mean explanation	Mean examples
	Stage:	Text	Ex.	Text	Ex.	Text	Ex.	Text	Ex.	Text	Ex.		
ICM (lazy learning meta-model)	Own	4	3	3		4	3	2	2	3	2	3.2	2.5
	Supervisor	4	4	3		4	1	1	3	3		3.0	2.7
	Expert	3	2	3		4	2	1	3	2	1	2.6	2.0
Softmax (numerical model output)	Own	3	3	3	3	3	4	3	2	4	3	3.2	3.0
	Supervisor	2	3	2		2	3	1	3	3	4	2.0	3.3
	Expert	3	3	1	1	2	3	1	1	1	3	1.6	2.2
Park et al. (active learning meta-model)	Own	2	3	3	3	3	4	3	2	4	3	3.0	3.0
	Supervisor	1	2	2		2	3	1	3			1.5	2.7
	Expert	1	3			1	1	2	1	2	1	1.5	1.5
Finds it useful:		Yes		No		Yes		Maybe		No			
Preferred measure:		ICM		Softmax		Park et al.		Park et al.		Park et al.			
Participant's explanation similar to:		ICM				ICM				ICM			

The table above shows the three sets of ratings (min. of 1 and max. of 4): 1) the participant's own belief of understanding (row 'own'), 2) the supervisor's belief and 3) the ML expert's opinion of how well the given explanations from the participant matches the measures and examples ("Ex." columns). The table also shows whether the participant found a certainty measure useful, their preferred measure in the end and the best match with their explanation of a certainty value. Blank values were not given or lacked clarity.

Both the experiment supervisor and the ML expert concluded that most participants had some degree of understanding for ICM. Only one participant was not able to comprehend the textual explanation but the understanding of ICM was on average rated higher than that of the 'numerical' and 'active learning meta-model' measures, by both the supervisor and ML expert.

The explanations about the numerical output were lacking because participants had trouble comprehending that a model can learn knowledge and represent it in parameters. They had less difficulty for ICM because its outputs related directly to past situations. The explanations from the participants regarding the 'meta-model' measure were the most inaccurate. Nearly all participants had the tendency to see this measure as combination of ICM and probabilistic output. This was also the reason why three out five participants tended to prefer this measure in the end, even though their own explanations of the certainty values resembled the approach used by lazy-learning meta-models.

2.5 IOSS prototype

IOSS is designed and intended to run on a tablet alongside the existing DP system, as this allows the DPO to roam the ship and take the system whit him. Furthermore, mobility can be facilitated even further by enabling smart watch functionality. On both devices, the main communication from system to user happens through visual and auditory signals. The embodiment (i.e., avatar) of IOSS consists of blue circles that slowly emanate from a centre point. (see figure 2-9).This appearance changes with respect to colour and pulsation speed in order to reflect the severity of the situation (e.g., rapidly emanating, orange circles in case of increased risk).

When the DPO and system first meet, the IOSS introduces itself to the user, explains its main functionalities and lets the user view and adjust settings with respect to the Smart Notifications. The goal of the introduction screens was to increase the predictability of the IOSS and therefore increase trust in the system (Johnson et al., 2014). Directability (i.e., enabling the user to set rules for the system), predictability, and trust building are key elements in order to establish an effective collegial relationship between system and user.

The first part of the introduction shows the two different types of Smart Notifications that the IOSS will send to the user (figure 2-9). First, it will relay DP alerts from the DP system. Second, it will make predictions about sensor warnings. In this design, a button was included which let the user scroll through several examples of these alerts and predictions, to gain insight in the possible Smart Notifications that the IOSS could send.

Figure 2-9. Introduction to IOSS.

After the introduction, the DPO was given the possibility to view the current settings for the triggers of the Smart Notifications (figure 2-10). The operator is free to adjust these settings based on the type of operation, the environmental conditions, or their own individual preference. The operator can set these working agreements (Bradshaw et al., 2004) by changing threshold values for triggering the Smart Notifications. For example, the user can choose to only be notified by the system if the wave height exceeds 2 meters, and the user is able to set how far ahead the DP error will be predicted.

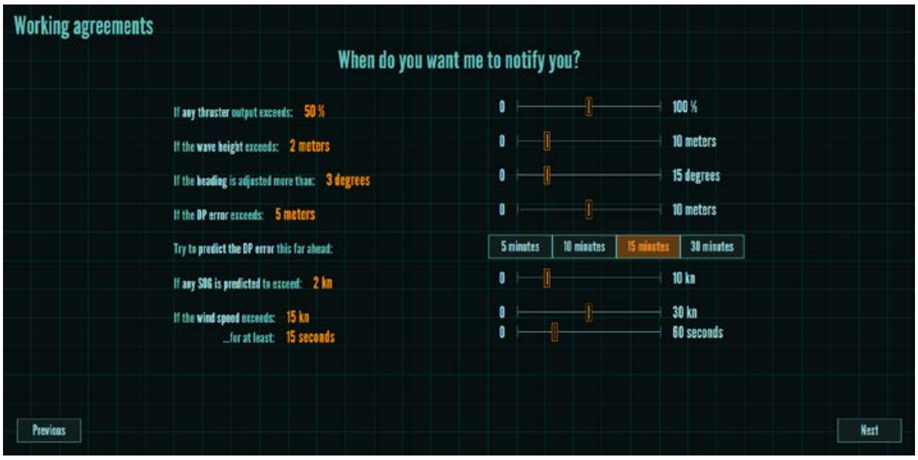


Figure 2-10. Creating working agreements

The adaptiveness of IOSS is further advanced by enabling the DPO to determine how tasks are divided between system and operator. That is, for each particular task the user may choose a level of automation for the system with respect to execution of a task (figure 2-11). For the prototype, a simplified list of three different levels of automation was selected, based on Sheridan and Human (1992) and Parasuraman, Sheridan, and Wickens (2000). By actively determining the working agreements and task division, the DPO is made partly responsible for the density of information that is communicated by the IOSS, and for creating and maintaining a proper balance in this information flow. Moreover, by explicitly stating the capabilities of the IOSS to the user, and by enabling customization of its communication policy, a foundation for trust between system and DPO can be established, and a better understanding of the functional differences between IOSS and the DP system is facilitated.

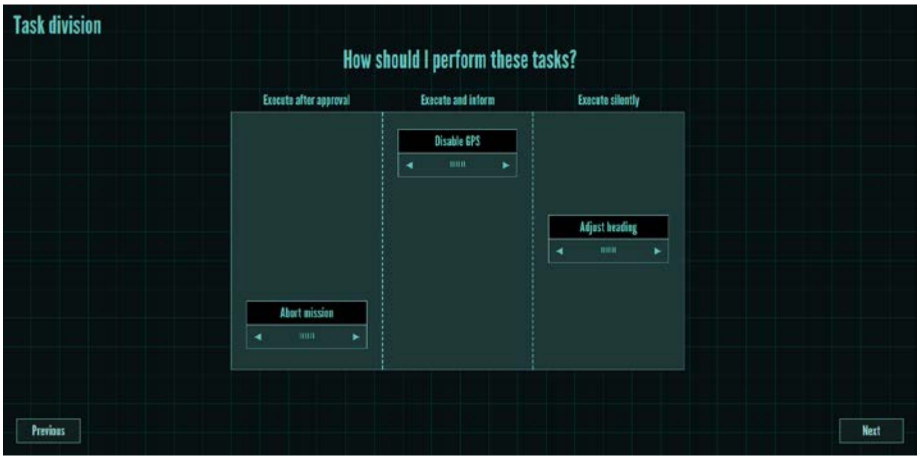


Figure 2-11. Determining the task division.

The main responsibility of the DPO is to monitor the draught, trim, and stability of the ship which is visualized by the DP system at the bridge, and to intervene and resolve (potentially) dangerous situations that are identified by the DP system. In this supervisory control task, the DP is at the bridge and the IOSS functions as a complementary monitoring system (figure 2-12).

Figure 2-12. IOSS (tablet on the left) as complementary system at the bridge.

In addition, a core functionality of IOSS is that it increases situational awareness by continuously processing and analysing relevant information from various sources (e.g., the ship, the DP system, weather data, wave data). This is a huge benefit as compared to the current (DP-only) system, because this system only provides information concerning the ship itself. All relevant information can directly be requested by the DPO at all times via the tablet-based GUI of IOSS. The main screen of this GUI only consists of the avatar that, by changing in colour and pulsation speed, allows the user to obtain a quick insight into the current situation. At any time, the DPO can obtain a more detailed status report by touching the avatar, which shows the menu as seen in figure 2-13.

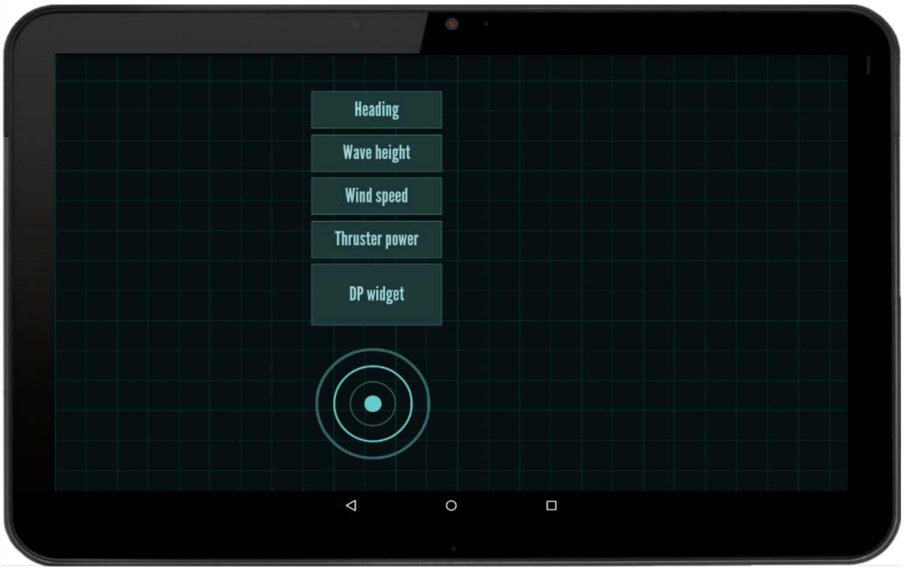


Figure 2-13. Main menu of IOSS.

By touching one of the menu buttons, a graph with real time information appears which illustrates the past and current status of the requested information (e.g., heading, wave height, or wind speed) (figure 2-14). Moreover, the DPO can directly obtain insight into the current status of the DP system by touching the ‘DP widget’ button, which shows a visualization of the current set- and centre-point of the ship, and the thruster direction and power (identical to the visualization in the DP system itself) (figure 2-15).

Figure 2-14. User-requested wind speed information.

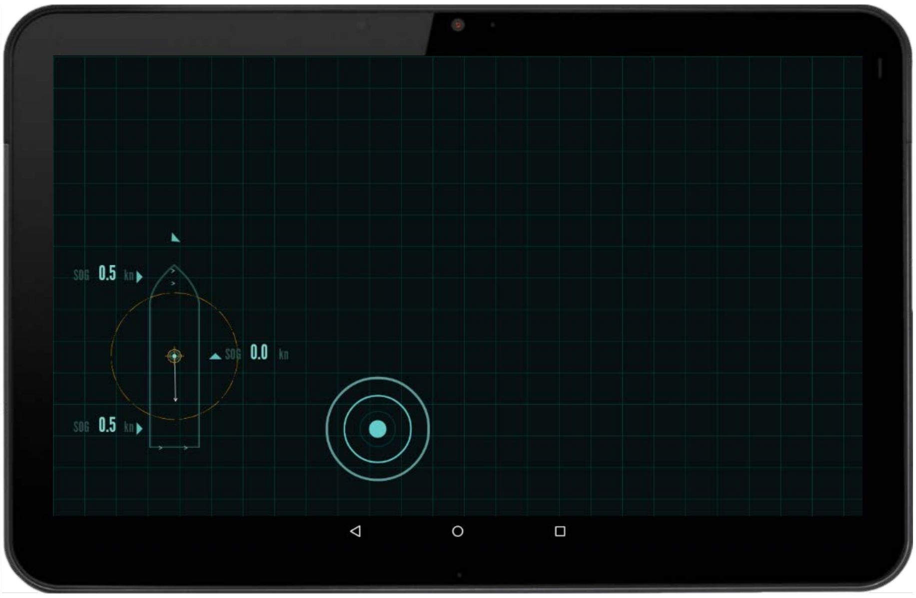


Figure 2-15. The DP widget of IOSS.

Another core functionality of IOSS is its predictive capability. That is, while the current DP-only system simply relays sensor data to the DP system GUI without running (predictive) analyses, IOSS continuously analyses the incoming data from various sources in order to predict the future situation, thereby increasing situational awareness. Dependent upon the working agreements that have been specified by the DPO, IOSS is able to provide personalized advice about the opportunity for the DPO to leave his/her workstation. Thus, while in the DP-only situation the DPO is obliged to maintain his position at the bridge in order to monitor the DP system, with IOSS (s)he is provided the opportunity to roam and engage in other tasks. An example of a such an advice is provided in figure 2-16.

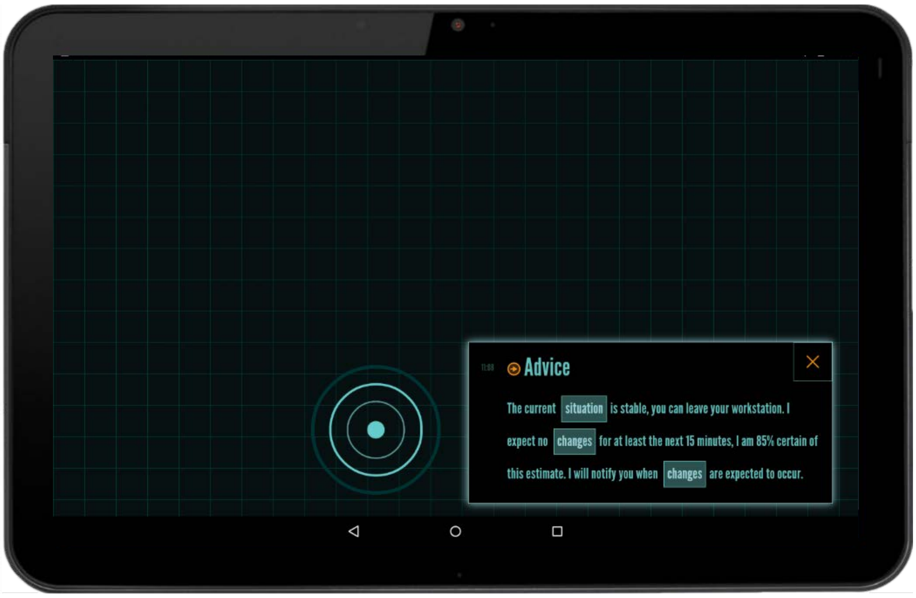


Figure 2-16. Advice to leave workstation.

In this example, IOSS notifies the user that it is possible to leave the workstation, because the current situation is stable and, most importantly, no changes in this situation are expected for at least the next 15 minutes. Moreover, this forecast is accompanied by an uncertainty margin, as calculated by the ICM algorithm (van der Waa, van Diggelen, Neerincx and Raaijmakers, 2018), which allows the operator the make better decisions concerning whether or not (s)he should leave the workstation. Explainability of the system is increased by enabling the user to press key words in the advice (e.g., ‘changes’ or ‘wind’), which shows a graph of the current and predicted changes of the information that is requested (figure 2-17). This is also important for building trust, because explaining the outcomes of the predictive analytics algorithms facilitates understanding by the DPO.

Figure 2-17. Advice to leave workstation accompanied by an uncertainty margin.

Based upon the advice and accompanying information that is provided by IOSS, the DPO can decide to leave the bridge in order to work on another task. In this roaming situation, the DPO can access the full IOSS system at any time on a tablet, and/or choose to take only the avatar along on a smart watch (figure 2-18). In the latter case, the DPO cannot access the visualized information from the sensors, but still has basic insight into the status of the current situation by the visual and auditory feedback that is provided by the avatar. In roaming condition, IOSS takes over the supervised control tasks of the DPO at the bridge, while keep informing and asking approval when necessary, according to the task division that was specified. Thus, while IOSS mainly takes over the monitoring and control task, the DPO can deal with other matters, while trusting the system to detect (un)expected changes in time, and notify the DPO accordingly.

Figure 2-18. IOSS on a tablet and smart watch.

An important functionality of IOSS is that it goes beyond a simple alarming function, by featuring smart notifications and adaptive system-user dialogue. For example, when the DPO is not at the bridge and IOSS predicts a situational change that needs attention, it advices the user to return to his workstation. By using proximity trackers, IOSS can adjust the timing of this message based upon the location of the DPO. Initially, the advice is concise and only provides a general and brief description of the situational change (e.g., changing weather circumstances), along with an uncertainty margin. By only presenting the most essential information on screen, the system prevents cognitive overload by the user. However, the DPO can choose to receive a more detailed overview of the situation by touching key words that refer to essential information in the message by the IOSS. For example, figure 2-19 shows additional information concerning the wind speed, after the user has pressed the word ‘conditions’ in the advice to return. This interactivity creates a functional system-user dialogue in which the user asks the system for a particular piece of information, and the system replies by providing the requested information in a visual and/or textual manner, accompanied by a one- or two-sentence explanation that also contains interactive keywords. The benefit of enabling the user to receive this cause-related information directly at his/her current position is that is prepares the DPO by stimulating the cognitive process at the workstation that is required to solve the problem. Thus, after (or while) receiving the advice (and gathering sufficient, additional information) by IOSS, the DPO returns to the workstation at which (s)he can directly resume the supervisory control task concerning the DP system.

Figure 2-19. Advice to return to workstation.

3 IOSS user evaluation

To verify the functionality and usefulness of the IOSS for the FPSO practice, it was evaluated by several experienced DPO's. In collaboration with a DP instructor, two FPSO specific scenarios were created for this purpose in which a failure mode was introduced at some point in time. Through a semi-structured interview the DPO's were asked for their opinion regarding the IOSS functionalities.

3.1 Participants

Five experienced DPO's were invited to participate in the study in order to evaluate the IOSS in a critical way. These operators were recruited through professional networks including those of the research partners, i.e. Bluewater B.V. and STC-group. The operators have operational experience in various DP domains, including but not limited to rock dumping, cable-laying, drilling, off-shore construction, and military operations for the Royal Netherlands Navy. The operators participated out of personal interest but received financial compensation for their time and travel expenses to the test location at TNO Soesterberg.

3.2 Scope of the evaluation

As explained in the introduction of this report, the IOSS was designed to fulfil specific operator support functions. In the evaluation study those functions were tested for their usability, the degree to which the operators found them useful and fit for future implementation for FPSO vessels.

The elements of the IOSS on which the evaluation was focussed will be elaborated in more detail in the following paragraphs. The first is where the IOSS introduces itself to the operator(3.2.1), then the notifications the IOSS gave were evaluated (3.4.5), usefulness of the various support functions (3.4.6), trust in the system (3.4.7), and finally an overall impression of the IOSS (3.5).

3.2.1 Introduction of IOSS to the operators

It is of critical importance that the operator knows what to expect from the digital assistant. To this end, the IOSS has been equipped with a few screens where it 'introduces' itself (see also chapter 2).

In the first introduction screen, the IOSS presented the information it could convey to the operator. This screen served to provide the operator with a good understanding of *what* the IOSS could be trusted to do. This included relaying information from the DP-system such as sensor values and alarms, as well as generating information itself using its artificial intelligence, e.g. expected implications of weather forecast.

The second introduction screen allowed the operator to make the so-called working agreements. In these working agreements the operator decided how and when the IOSS had to get involved in the DP task. This screen served to provide the operator with an understanding of exactly *when* the IOSS could be expected to give notifications.

Finally, the operator could make agreements with the IOSS regarding *who* would get decision authority between the operator and the IOSS. For example, simple tasks could be delegated to the IOSS so it could make executive decisions independently. Contrary the operator could also shift the task allocation away from the IOSS and remain in executive control. The question here was first; whether the operator would trust the IOSS to decide on, and execute, actions independently. Secondly, whether the criticality of the action decided on would influence the trust in the system (e.g. switching GPS antennae vs. aborting the mission). Additionally, the manner of allocating tasks between the operator and the IOSS was of interest, e.g. did the operator prefer to choose one task at a time, or decide for clusters of tasks at once

3.2.2 Notifications

IOSS conveys its information through conversation-style notifications (see design specifications, paragraph 2.3.1). Of interest was whether these notifications were understandable to the operator. As the concept involves the IOSS supporting the operator while roaming, it is of great importance that the notifications the IOSS gives hold enough information for the operator to maintain an adequate level of Situation Awareness. Therefore, the operators were questioned regarding their understanding of the notification.

The notifications themselves were also evaluated regarding the Intuitive Certainty Measure (ICM, described in paragraph 2.4.2 in terms of how useful this was to the operator, and also how this impacted the level of trust in the IOSS and its suggestions. Furthermore the semantics of the ICM were varied according to the design specifications and were also evaluated. Operator preference was evaluated regarding the conveying of the ICM.

3.2.3 Use of the IOSS during roaming

During the scenarios the operators were observed regarding the way they made use of the IOSS while they were roaming, i.e. away from the DP desk. Of interest was whether or not operators would feel inclined to regularly check the IOSS, or whether they would be comfortable enough to trust the system to alert them on time. This all comes down to the level of trust they had in the system.

3.3 Evaluation design

The operators participating in our evaluation were invited to our research facility. The evaluation was done with one operator at a time, and took roughly 4 hours per person.

3.3.1 Setup

The setup consisted of various systems linked together. One monitor was placed in front of the operator, linked to a computer running the Dynamic Positioning system, developed by RH marine (Conning 4500). This computer also served as a server platform for the IOSS system which was tied into the RH marine DP system (today called RH Marine Group). This same computer also ran the simulations and fed this information into the DP system and the IOSS.

A tablet computer was used to display the IOSS (web-interface). The operators could take this tablet with them while roaming and browse through the available information.

Figure 3-1 Impression of the IOSS in operation while roaming.

3.3.2 Scenarios

To demonstrate the functionality of the IOSS two scenarios were created and executed in a simulated DP operation. Both scenarios started off with a stationary DP operation in calm weather. After a period of 10 minutes of active monitoring the DPO was informed by the IOSS that it was safe to leave the DP desk for at least the next 15 minutes, and leave the supervision of the DP system to the IOSS. At this point the DPO's were taken away from the DP system and given access to the IOSS through a tablet computer.

During this period they were allowed to use the IOSS to maintain their level of SA, but were not instructed to do so. After a period of roaming a failure mode would occur during the scenario. In the first scenario this entailed a sudden change in the external environment, i.e. a sudden increase in windspeed. In the second scenario this was an issue related to the technical state of the ship, i.e. a thruster failure.

When the failure occurred, or was expected to occur in case of the windspeed increase, the IOSS would inform the operator and the operator had to respond to the situation. Both scenarios started out with a stationary operation during calm seas, this was necessary to allow the operator to go roaming for the purpose of our evaluation.

3.3.3 Semi-structured interview

A semi-structured interview was used to question the operators. This method was chosen to allow the operators to provide input at their own initiative as well.

The interview consisted of several parts corresponding to the phases in the experiment. A copy of the semi-structured questionnaire with the original feedback is included as appendix A.

During the introduction of the IOSS questions were asked after each of the three screens (introduction, working agreements, and task division). Questions related to the extent to which the introduction screens were clear in conveying messages. This was determined by engaging in a conversation with the operators and encouraging them to explain in their own words what they thought the screens actually conveyed and comparing that to the intent of the designers.

Related to the working agreements, the operators were asked how they would like to choose the setpoints for being alerted by the IOSS, and how far in advance they would like to be warned.

For the manner of dividing tasks the operators were asked how they interpreted the three variations of task division (i.e. “execute after approved”, “execute and inform”, “execute silently”). Additionally they were asked about how they would like to choose between these three levels (e.g. choose for one task at a time, or multiple at once) and whether they would want to have an additional level.

Following the introduction of IOSS the DP simulations started. The questions that were asked during the scenarios related to the notifications that were given by the IOSS to support to operator. The operators were asked about the extent to which the notifications were understandable, whether they would give follow-up to the notifications, and regarding the added value of the ICM. Furthermore they were asked if they felt that information was missing in the notifications, and to what extent they felt that it was still necessary for them to be involved in monitoring the DP system.

Following the execution of the DP simulations and answering the related questions, the operators were asked in more detail about the workings of the ICM. The certainty of the IOSS when giving advice can be calculated using several methods. These methods and their evaluation are discussed in detail in paragraph 2.4.2.1.

3.4 Results

3.4.1 General impressions

The operators’ reactions to the concept of the IOSS and roaming were positive. Multiple operators explained that they saw “the added value” and that it “definitely” had potential. One of the operators said that

> *“this would have made my job a lot more fun”.*

They recognized that it would be interesting for companies “to save on people, to save on money”, but more importantly, they saw opportunities for DP operators to be able to perform other tasks instead of monitoring the DP system. Monitoring DP was time “that could be spent in a more useful way” and it would made the DP operator’s job “more enjoyable”.

However, multiple operators added conditions to these positive expectations. Many underlined the importance of reliability, adding that such a system would be helpful “in an ideal world” and that it “has to be really watertight”. They noted that such a change would have to be made in close cooperation with the industry and would also have consequences for training of new operators: “you have to be able to expect that the person sitting in front of this, interprets the information correctly”.

Most of the operators were hesitant to the idea of implementing such a concept for every type of operation, but were open to the idea in the case of “stable operations” where the risks were lower. One of the operators already noted the ability to retain some form of control over the system: “a system can take over more tasks, provided you are able to set the thresholds”.

While operators saw the opportunities of being able to freely move around instead of having to stay in the immediate area of the DP desk, they also underlined that currently it is not acceptable at all for DP operators to leave. As the operators put it, “a DP operator cannot leave his station, always has to be on the bridge“ and “I assume that, at the moment you activate this, you stay on the bridge. You stay with this device. You cannot leave the bridge as a DPO”. One of them even said that “as a captain I won’t accept that, that they would walk away”.

Some of them were reluctant about the idea of leaving the room. For them, leaving the console meant that they were free to move around on the bridge, not necessarily moving to other parts of the ship:

> *“If that works it would be very useful. Not to really leave, but to do something else in the same room, or an adjacent room.”.*

The participants also noted that even when an operator stays physically near the DP desk, his attention could shift once he starts performing other tasks: “the moment you start doing something very different and you focus on that, it could mean that [DP] goes to the background”.

3.4.2 IOSS Introduction screen

The first of the three opening screens shown to the operators was a mockup of the “Introduction” screen. The screen introduces the IOSS and its avatar, and describes the two types of notifications that the IOSS would send: notifications based on relaying DP alert from the DP system, and notifications with predictions.

The goal of this screen and the difference between these two types of notifications was clear to the operators. As one participant put it, it meant “that I can receive a notification when I walk away from the desk, or a prediction”. Another said “this is all clear, they are also logical things to receive”.

According to the operators, the notifications showing DP alerts “should be recognizable” to “everyone who has worked with a DP system before”. However, the notifications showing the predictions raised more questions. While the operators saw the benefits of such notifications, they were curious as to the sources of the (weather) data, “Where do you get this prediction from? How does it know the weather?”. According to one operator, it would be good if it was “fed with up-to-date data” from online sources.

Again, the operators were positive but cautious, saying that predicting the weather is “tricky” and “really hard”, and that “if the weather is a bit critical, I would be careful” but “if the weather is stable, then it would probably be easier [for IOSS] to make a prediction”. If risks increased, they would not rely solely on the IOSS: “What I imagine, is that if you apply this, it would be during favorable weather conditions. And that if this device tells you the conditions are bad, that you are going to pay attention anyway.”

One of the operators reasoned about the types of predictions the IOSS would be able to make: “What it can’t predict, of course, is a thruster failure or something like that”. Another informed about the possibility of setting thresholds holds, asking “Can you set values yourself, for predictions, when you get a warning?”

3.4.3 Working agreements

The ability to set such thresholds for when the IOSS would send a notification, was shown in the second mockup, titled “Working agreements”.

At first, the operators were only shown the mockup, without any additional explanation. However, the concept and the use of making working agreements was clear to the operators. One of the participants immediately said, “ah, [this is] when you are informed” and another called them “alarm conditions”. One of them said that

> *“being able to assign when it’s going to give which notification yourself, is essential, I think”.*

The ability to drag the bars on the screen to set thresholds was obvious to some, while others first had to interact with the system to discover this feature, which points to a lack of clear affordances.

The operators had suggestions for additional settings, namely separate thrusters, GPS-related values, the available power and the water current. Wave height, on the other hand, was not deemed necessary “for large ships”, also “because it has a lot to do with the wind”. Some of the settings could be renamed. For instance, DP error could suggest that there was something wrong with the DP system and could instead be named “position offset” or “set point error”. Other settings, such as the heading and DP error, should have smaller, more precise values.

There were different opinions on who should have the ability to set these working agreements. Some participants thought the working agreements should be set by the DPO, in order to adjust them to the situation at hand: “these are things that are dependent on the situation”. A downside was, according to the operator, that “those things would become part of your hand-over”. Therefore, the number of settings should not be too high: “then I have to transfer them to you, and then you have to verify them with your checklist, so you shouldn’t have too many to hand over”. Asked if such working agreements were useful, another operator put it this way: “Yes, but not too many. Those are things that will be changed once and then are overlooked.”

Another opinion was that the default settings, or the setting parameters, should be set by the commander or captain. One operator said: “I would not be a proponent of individual profiles. […] I can imagine that if I were captain, I would want a certain influence over this, so people cannot switch of everything. Because you have to be this watchdog.”. Another said: “I see potential in profiles for working agreements. But like, there is a value which I cannot exceed or go below. This max/min values could be set with the commander and for instance be password-protected."

Yet others said that the settings should be set based on the nature of the current operation, “I would pre-program them”, based on the environment, “profiles based on location, for instance always have this profile on open sea”, or based on the ship, “I see working agreements as something you set once and then never again, really specifically for our ship but I would not want to set these again every time”.

Multiple operators asked for more complex or smarter thresholds. For instance to take into account averages and the rate of change over time. One operator would have liked to be able to set a limit for “the change of the wind direction in degrees, for instance 30 degrees, opposed to the average over the last hour”. Others asked for time limits instead of simple limits: “Maybe you could add a time scale. So that if you hit the 25 percent once, there’s no big deal, but if it’s there constantly for a few minutes, or if it hits it a few times, then it activates”. Another explained: “Wind speed is hard. Because nobody knows how long a gust of wind will keep. If the wind speed suddenly increases, then it’s hard to say whether you want a warning. But you would want one if it happens a few times over a period of time”.

One of the operators asked for “different gradations” of notifications, for instance warnings and alarms, depending on the impact.

After going through the first and second scenario with the demonstrator, the operators were asked if they would like the option to revisit their working agreements. All of them said they would have liked to be able to revisit this screen, although not everyone would have changed the original settings. That would only be necessary if the conditions of the operation had changed: “If the work stays the same, I would not adjust these very quickly. Look, if the work changes, then you should have another look.”

3.4.4 Task division

The third and final introductory screen showed the Task Division feature. This would allow operator to choose the level of automation for specific tasks. In this mockup, three tasks were included that could be set to “Execute after approval”, “Execute and inform” or “Execute silently”.

After interaction with the mockup, the operators quickly discovered how to use this function. The three different levels of automation were also understood rapidly. Although the operators asked questions to confirm their impressions, they all had a correct understanding of what the levels entailed.

The operators were wary about choosing “Execute silently” option for tasks:

> *“I think this one [Execute silently] is tricky, that’s really letting go of a large bit of control. Even if this column would be available, I don’t know if I would use it.”.*

Other operators said:

> *“I think Execute silently would be rarely or never used”*

and

> *"But execute silently, that's… Then you have nobody to take it up with, right. If something happens that it's: 'yeah, that was the computer'. At least with execute after approval your operator is in the know."*

Only one of the operators was comfortable with choosing this option, although again it was dependent on the operation and the specific task: "A number of things he can definitely do automatically, as long as it doesn't endanger safety."

On the other hand, multiple operators liked the ability to choose between the first two levels, namely "Execute after approval" and "Execute and inform": "Nice to be able to move this around, this gives you different possibilities for different situations. […] But yes, indeed, inform". Another operator said "Almost all actions can be initiated, but do have to have an approval". When asked about the fact that many tasks are already automated and performed by the DP system, the response was: "Sure, but you're still there. And you see what's happening. I would find it really odd if you would walk away and the ship would start turning."

The operators were also asked which settings they would choose for these tasks and what their approach would be. Again, the operators were cautious. Most of the participants said that they would start by putting all tasks in the left-most column, "Execute after approval" and then move them one by one to one of the columns with a higher level of automation, depending on the task. For instance, one operator said he would "just have a single reference point, and then add exceptions". Another said: "I think that by default, everything will be in the left column. And then move this and that to the middle column. […] I think that is the safest option."

What they agreed on was that it was useful to at least go over these settings, "that you select them yourself every time you start the DP" because, as they put it, "it also gives you a moment to consider it". The operators also agreed that this process should be repeated for every task, and that if they should be set, it should happen one by one.

Although the operators were cautious about letting go of manual control of the system, multiple operators said that over time, they expected to see a shift from manual operation to higher levels of automation: "As you work longer with the system, you will see a shift to 'Execute and inform'". This would also depend on operators' experiences with the system: "But if things often go well, and it is a reliable system, that grows of course, that you move to 'Execute and inform'"

3.4.5 Notifications

For each scenario and each notification questions were asked separately.
There was some overlap between the questions, therefore the following paragraphs provide an aggregated account of the comments that were made by the operators based on the research questions. Some comments were made regarding highly domain-specific improvements, for the purpose of evaluating the functionality of IOSS these were excluded. Instead the focus is on the design of the IOSS, and the manner of conveying information to the user.

3.4.5.1 Information presentation

In general all operators indicated that they thought the notification were easy to understand. There were no misconceptions as to their meaning or intent. However there was some feedback regarding the manner of presenting information. It was noted that it would aid the understanding of the operator if graphs displaying information used a fixed scale, and display trend data going back at least several minutes. Some operators also indicated that they would like to be able to see the variance of certain sensor values. Another operator also suggested that instead of displaying all relevant sensor-values, it could be useful to summarise these by using something like a checkbox or a traffic light. If all systems are nominal the light would be green, and if there are issues it will turn to orange or red. The operator felt this could aid the speed of interpreting the situation, and is in correspondence with colour-coding currently used in DP operations.

3.4.5.2 Phrasing

One important piece of feedback we received is to mind the formulation between different warnings. The first notification the operators got was to inform them that it was safe to leave their desk [*The current situation is stable, you can leave your desk...*]. At the end of the scenario when the weather had become unstable, they received a notification to remain at the DP desk [*The current weather conditions will remain the same … advise you to stay…*]. Some operators felt that these sentences were too similar, and whereas the intended meaning of the latter notification was to inform the operator that the situation would continue to be unstable, it was interpreted as no changes being expected.

3.4.5.3 Providing follow-up to notifications

Providing follow-up to a notification gives an impression of the extent to which the operator took the notifications seriously and found them useful. While the notifications were generally easy to understand, there was still some difference between the operators when they were asked about the degree to which they would give follow-up to the message.

“Not until I trust the system”

“Yes! Because you can see all relevant information, but I would still check regularly until I have more experience and trust the system”

“Maybe I wouldn’t stay behind the desk, but I would stay in the area”

This selection of thoughts is a reflection of the general feeling about the IOSS. Most operators felt comfortable enough to follow-up on the notification, however they felt a need to regularly check the system and confirm whether it is still functioning. It was mentioned by most that more experience with the system would allow it to gain their trust and they could see themselves rely more on the system.

Interestingly though, during the simulated DP operation, when the operators were allowed to go roaming, most actually did not regularly check the IOSS to confirm whether it was still working. It is not entirely clear what causes this discrepancy.

It might be because the scenarios began relatively calmly and nothing suggested a cause for concern, or it might be that trust in the system develops quicker than the operators themselves might have anticipated.

3.4.5.4 Usefulness of certainty measure

An important element of the IOSS is the Intuitive Certainty Measure. While its workings have been discussed already in paragraph 2.4.2, the operators were asked after each notification including an ICM, what they felt about the IOSS providing an indication of its certainty. The most commonly heard sentiment was initially that the operators felt no need to be informed regarding the certainty of the advice by the IOSS. However, as the scenario progressed and they were confronted with the fact that the IOSS could make mistakes, the operators believed a certain certainty indication would be very important to maintain trust in the system following uncertain prediction. In short, the IOSS was allowed to make mistakes, provided it alerts the operator in advance about the likelihood of a mistake occurring.

3.4.5.5 Quality of information provided by IOSS

The operators were asked about the content of the notifications. This included questions regarding the amount of information, the presentation of the information and whether information was missing. Again, the focus is not on highly domain-specific comments, but rather those focussed on the workings of a digital operator assistant.

One of the comments that was given was related to the certainty of the predictions the IOSS gave. Some operators indicated that they felt information was missing regarding the cause of uncertainty in predictions. They felt it was not sufficient to simply say the certainty is low, but they want to know why. Similarly, when the IOSS mentions for example that an increase in wind is expected, the operators want to know exactly what is expected and how this expectation came about.

> *"I want to know where the uncertainty comes from, and I want the system to summarise whether everything is under control or not."*

This quote was discussed in some more detail with the operator, who made an analogy between the IOSS and a human crewmember. The operator said that he is aware that humans can never be completely sure about something, which is accepted, however he relies on their judgement to say whether it's under control or not. The conclusion of this statement is that the operator felt it is fine that IOSS is uncertain about predictions, as long as it can be trusted to monitor the system correctly and adjust its predictions when the situation changes.

Regarding the presentation of the information a comment that was made more than once was related to the manner information was requested from IOSS. In the current version each system, i.e. wind sensor, thruster output, etc., had a unique button for requesting information. Some operators felt it would be more convenient to have just one 'more info' button, that would display all relevant information. Care should be given here to ensuring the proposed 'more info' button does not result in making a redesigned copy of the DP system, but should only provide information that is essential within a certain context.

Another interesting remark that was made by some of the operators was following a notification saying the situation would remain stable. The question was whether the operators would feel a need to monitor the system through the IOSS while roaming after that notification.

“No, if the IOSS says it will remain the same, there is no need for me to look.”

This indicates that once a certain level of trust in the system is achieved, operators would feel quite comfortable relying on the IOSS, and trusting it to notify them on time as soon as the situation is expected to become unstable

3.4.6 Usefulness of IOSS functions

Towards the end of the evaluation, the operators were asked to rate each of the functions that were available as useful or not. Table 3-1, provides the results of this exercise.

Table 3-1. Usefulness of IOSS functions.

Function	Pp1	Pp2	Pp3	Pp4	Pp5
DP status screen.	V	X	V	V	V
The possibility to request more information through the notifications.	V	X	X	V	X
Notifications of DP Alarms	V	V	V	V	V
Notifications about being able to leave, or return to, the DP station.	V	X	V	V	V
The graphs in the smart notifications.	V	V (I want to see the history over the last few minutes)	V (only with a constant scale)	V	X
The certainty of the advice by IOSS being correct.	V (if you know the reasoning behind it, I don’t need to see the %)	V	V	V	V (but represented more simple)
Explanation of how IOSS concluded that you could leave or had to return to the DP station.	V	V	V	X	V
The possibility to make more agreements with the IOSS about which notifications are given when.	X	X	X	X	X
The checklists that IOSS can fill out depending on which information you have already seen or not	X		X	X	X

Function	Pp1	Pp2	Pp3	Pp4	Pp5
seen, and which information has changed.					
The ability of the IOSS to detect whether you are near the DP station.	X	X	V (important)		X

3.4.7 Trust in system

It has already been mentioned in the previously discussed comments that trust in the system is a recurring theme in the assessment of the IOSS by the operators. It was mentioned as being important in deciding whether or not to adopt the suggestions made by the IOSS. Additionally, most operators mentioned that they want to be able to check how and why the IOSS came to certain predictions. This could be a sign of initial lack of trust in the system, that might subside after the operators have had more experience with a system like the IOSS, and build trust in its capabilities. Interestingly it was found that during the DP simulations, the operators rarely checked the IOSS while they were roaming. Even though they indicated that it is important to be able to do so, most of them didn't out of their own initiative. This could be a result of using a simulated task in a controlled environment without any risks, but it provides a promising outlook on the employability of the IOSS.

The operators were also asked how the quality of advice would impact their level of trust in the system. The example as provided that IOSS would give them an advice that would turn out to be incorrect. While all operators understandably agreed that wrong advice would decrease their trust, interestingly enough they mentioned that their trust in the system would not be impacted, as long as the IOSS mentioned it was uncertain at the time it gave the wrong advice. Despite the varying opinions about the ICM, it seems that it nevertheless can serve as a great mediating factor for preserving trust in the system when certainty is indeed low.

3.5 Overall impression and conclusions

At the end of the session, the operators were asked to grade the IOSS on a scale of 1 to 10. Although two of the operators did not express their opinion in a grade, two others gave the IOSS an 8 out of 10. The fifth participant took the type of operation into account, and graded it a 9 for "stable operations" and a "6 or 7 for more complex rock dumping". The two operators that graded it with an 8 also attached conditions to this grade. One of them said:

> *"[it is] dependent on the type of work, I think it will not be possible to use it everywhere, but if the work allows, that there's an added value."*

The other explained

> *"It seems useful to me, but we work very differently. By that I mean that you must always stay on the bridge. Business management is an obstacle but I do think it's useful. I'm giving an 8, with the comment that a lot will have to be changed in the regulations."*

These grades reflect the general conclusion that can be drawn from the interviews: a positive reaction to the IOSS concept, but the success depends on the situation and on the adaptation of regulations and business processes.

It also reflects a tendency to want to remain in control of the system, as can be observed from the comments about the Working Agreements and the Task Division. Although the operators saw the opportunity of higher levels of automation, they were not yet convinced that the system would be reliable enough. Therefore they appreciated the possibilities to customise the system to their own judgment.

Overall, most of the operators were positive about the certainty measure, saying that it tells them what to expect and gives them more insight.

On the other hand, there was also some doubt regarding some of the functionalities of the IOSS. Multiple operators said that “you have to start somewhere” ….

The operators also identified several risks for the IOSS. The first was that, even if the IOSS helps to monitor the system, this could still result in information overload: “I think you have to watch out that you don’t again have too much information. That this [the IOSS] becomes a full-time job again.”. A second was that although it would be helpful if the IOSS was aware of the location of the operator, there was a definite risk of invasion of privacy: “Monitoring too much is not that ethical anymore”.

In conclusion, the concept of the IOSS as presented to the operators was received well. There was a general belief that a system like the IOSS could be of great benefit during DP-like operations. Some remarks were made that indicate the importance of human-machine trust for a successful collaboration to become possible. While the operators believed that the IOSS provided good information, a need was felt to double check the suggestions given by the IOSS, and compare those to how the operator would judge the situation himself. This reflects the level of trust of the operator in the system. It is important to ensure an appropriate level of trust to avoid automation disuse, and prevent automation misuse (Parasuraman & Riley, 1997). Ultimately the goal should be that the human knows in which situations the system is reliable. This can be achieved through operator experience, but also though assistance of the ICM.

4 Adaptive Support for Supervisory Control Operators in Highly Reliable Automated Systems

Semi-autonomous system performance is highly dependent on how operator and automation function as a team. Both components are highly interdependent. Between manual control and full automation, different levels of automation, or collaboration forms, can be distinguished. Well known classifications are made by Sheridan and Verplank (1978) and by Endsley and Kaber (1999), with different variations, but others exist. A special form of human-automation collaboration are adaptive systems. Adaptive systems are systems in which the locus of control varies over time. This implies that the responsibility for a specific subtask moves from the automation to the operator or vice versa. Adaptive automation is a subset of adaptive systems, implying that the automation takes over tasks from the human operator when the need therefore arises. Adaptive automation has been mostly utilized in situations of underperformance due to high operator workload, for instance in fighter jets.

4.1 The present research

In the present research we investigate whether adaptive automation could be used to address operator's cognitive underload states in highly reliable automated systems and potential out-of-the-loop (OOTL) performance problems, specifically loss of SA. This has not yet been investigated before. More specifically, we investigate the potential of an adaptive algorithm that assesses whether the operator is paying attention (or not) and decides, based on the situation and hence, the need for operator involvement, whether the operator needs to be alerted. The use of such a hybrid triggering strategy using performance measurement to determine whether an operator is paying attention (or not) is also not studied before (see also Kaber, 2013).

The advantages for Dynamic Positioning (DP) operations are obvious, justifying this research for practice. Traditional alarming systems support the DP-operator by showing alerts and alarms when pre-set values are exceeded. Still, the operator has to monitor the system constantly to detect deviations on time. This makes it quite impossible to perform other tasks concurrently. Adaptive support could allow the operator to perform other duties, without losing overall awareness of system state. This could alleviate the consequences of cognitive underload during periods of boredom, when the system is working perfectly, while at the same time it could help the operator to stay 'in the loop' while concurrently performing other duties on board of DP operated vessels. In the remainder of this chapter we describe the research questions and hypotheses underlying this research effort. Then we present the design of the research, the results of the experiment and the conclusions.

4.1.1 Research questions

The present research is set up to investigate whether adaptive automation helps in achieving higher overall system efficiency, as compared to situations without adaptive automation (i.e. no adaptive algorithm).

We are also interested to see whether this adaptive support could provide for additional benefits, as compared to static (leave on) support, e.g., in performance or reduced workload (cf. Parasuraman et al., 2009).

When in the face of a threat to the operation the DP system malfunctions, the operator has to act swiftly and accurately. The environment in which the DP operator has to perform is sometimes challenging with distractions and environmental stressors present, such as sleep derivations, time pressure, danger or loud noises from ongoing drilling or pipe laying operations. These conditions could harm supervisory control. In the present research we investigate whether this is true. We want to see what the effects of stress are on supervisory control and whether support could alleviate some of the negative effects of stress.

4.1.2 Hypotheses
Based on our research questions we formulated the following hypotheses:
- H1. We predict that performance and situation awareness would be enhanced with adaptive support, whereas overall mental workload would be reduced, and that these benefits would be specifically associated with adaptive (as opposed to static) support.
- H2. We predict that stress has a negative effect on performance and situation awareness.
- H3. We predict that the effects of adaptive support are more pronounced under high levels of stress as compared to low levels of stress.

4.2 Methods & Materials

4.2.1 Participants & design
A total of twenty-six students took part in this study. One participant was excluded due to a failure of the computer to record the data. The data reported here is based on the remaining twenty-five participants (13 male, 12 female). Their age ranged from 19 to 44 years (M = 27.5, SD = 7.8). Each participant had normal hearing and vision. None had prior experience with dynamic positioning but some had participated in a precursor study. Participants were paid €45 for participation and were promised another €40 for the best performing and €20 for the second best performing participant on the task to enhance motivation.

We tested our hypotheses in a 2 factor mixed design. Support was manipulated within subjects and had 4 levels: no support, static support, adaptive support based on critical events; and adaptive support invoked by a hybrid triggering strategy. Stress was manipulated between subjects and had 2 levels: high stress and low stress. Both experimental conditions are explained in more detail below.

4.3 Experimental manipulations

4.3.1 Adaptive automation
The operators in our study are aided by a predictive warning system, which provides support (Level Of Automation: LOA) at four levels in terms of Sheridan and Verplank's model (1978):

- LOA1, the operator is required to find emergency events manually, receiving no support from the system;
- LOA2, the system alerts the operator when system values change, but the operator is required to carry out emergency management manually;
- LOA3, the system alerts the operator when system values have significant changes, and an emergency event is likely to occur, but the operator is required to carry out emergency management manually;
- LOA4, the system informs the operator when an emergency event is likely to occur, only when the operator is not paying attention, but the operator is required to carry out emergency management manually. Hence, a hybrid triggering strategy is utilized for invoking adaptive support, based on (a) critical events (i.e. surpassing thresholds and significant rise or drop of relevant system values) and (b) cognitive state (is the operator paying attention to critical events?). An eye tracker was used to infer whether participants were paying attention to the relevant system values or not.

4.3.2 Stress
Noise was employed as a stressor, representative of a class of ‘arousal-inducing’ stressors (similar to time pressure, danger, heat, etc.) (Sauer et al., 2013). We used a low and a high stress condition. For the high stress condition, industrial noise was administered through headphones by a digital recording, comprising different types of machine noise (drills, power saws, etc.). It was the same recording that was used by Sauer et al., (2011). Noise level was set to 75 dB(A). Based on guidelines for maximum exposure times as a function of sound pressure level (Meyer-Bisch 2005), a level of 75 dB(A) was considered sufficiently loud to be perceived as a stressor while not endangering the participants’ hearing or health. Prior to the beginning of the experiment, participants were informed that they would be exposed to noise during the testing session. For the low stress condition, pink noise was administered at 50 dB(A). Both conditions were extensively tested by sound engineers and approved by our institute’s ethical committee.

4.3.3 Experimental task
We developed an in-house simulation capability, the DP simulation, designed to isolate some of the cognitive requirements associated with a single operator controlling maritime vessels operating on Dynamic Positioning (Barnes et al., 2006; Van der Kleij et al., 2018). The goal was to create a microworld with face validity for future DP operations, while providing a degree of experimental control. The design of the simulation was based on real DP systems and actual sensor information. Scenarios were based on a critical events analysis. When a critical event occurred, participants had to decide on the best action to take. When a decision was made, the scenario stopped. The task was trained to criterion during training sessions and all required information for a decision was made available during the experiment, to relieve working memory load. In figure 4-1, the experimental setup is depicted.

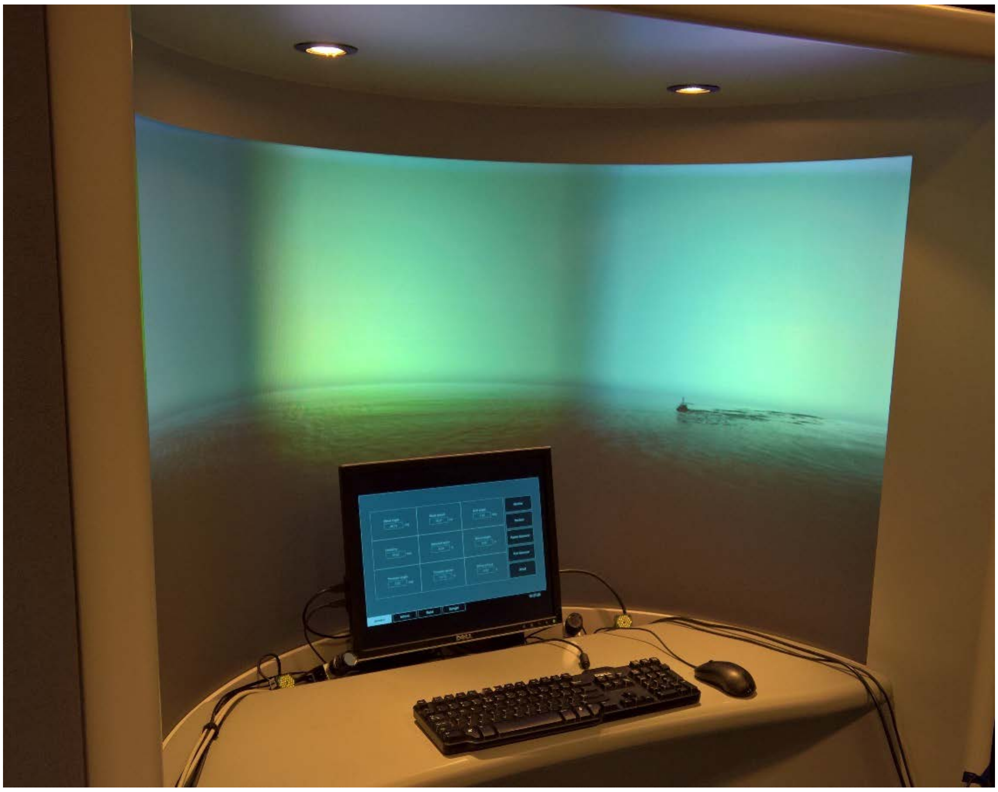


Figure 4-1. Experimental setup, showing the workstation, eye tracker equipment on both sides of the computer screen, and a simulated outside sea view.

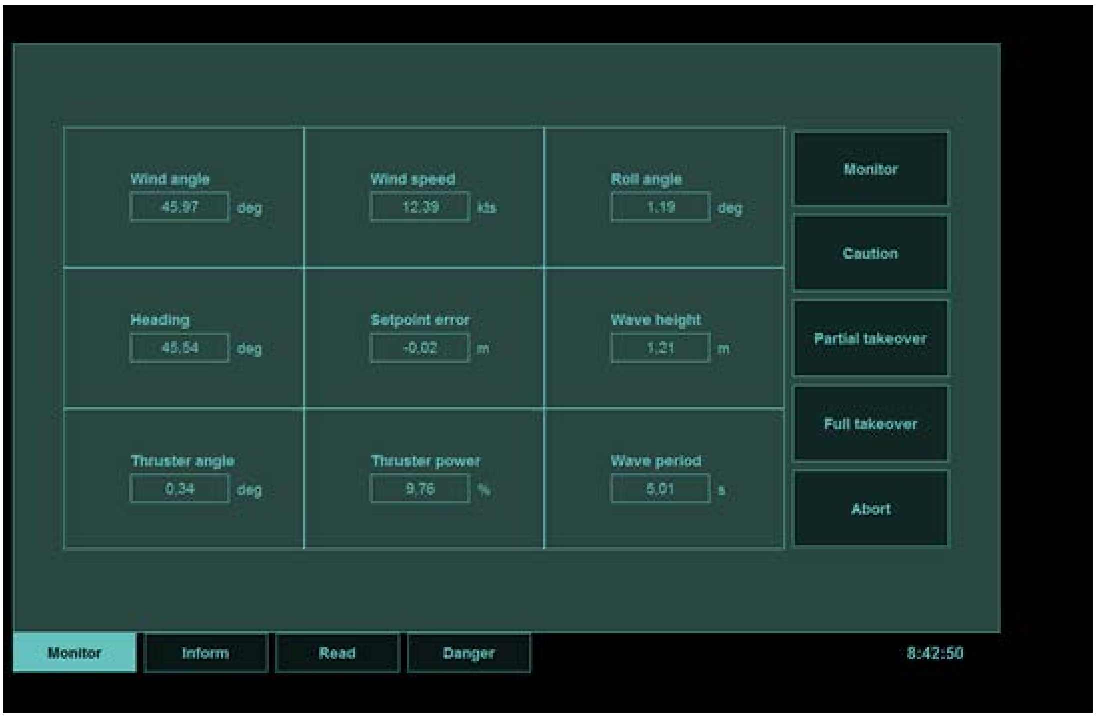


Figure 4-2. DP user interface showing important sensor information and action buttons.

To make the DP task more realistic and to be able to measure secondary task performance, several tasks were added to the task simulation (see figure 4-3, figure 4-4,and figure 4-5). The operator was required to complete four tasks, which were defined as primary and secondary tasks. The primary task was to monitor the safe operation of the automated DP system and react to dangerous situations (see figure 4-2). The secondary tasks were: (a) to inform management of values of sensors at regular time intervals (i.e. at 1-min), measuring prospective memory performance (see figure 4-3); (b) An administrator task (self-paced) marking specific predetermined words in a text file (see figure 4-4); and (c) to detect and report threats (incoming ships) at sea, visible on the simulated outside view (see figure 4-5). One ship would appear at a time. All ships started their approach from the same distance which was invisible to the operator. They approached at equal speeds but from varying angles, between -70 to +70 degrees.

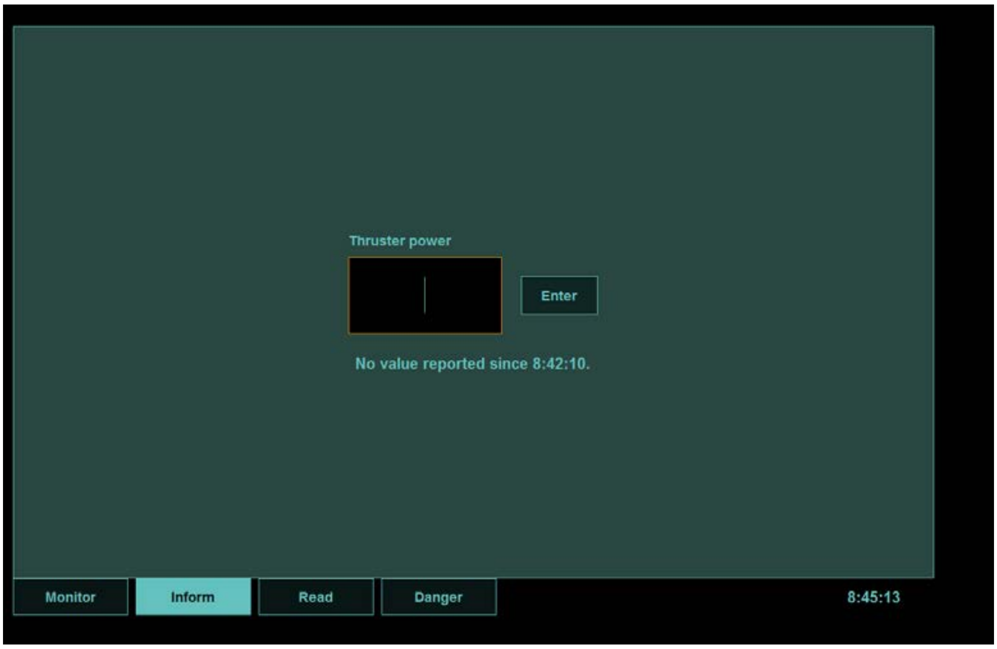


Figure 4-3. Inform secondary task.

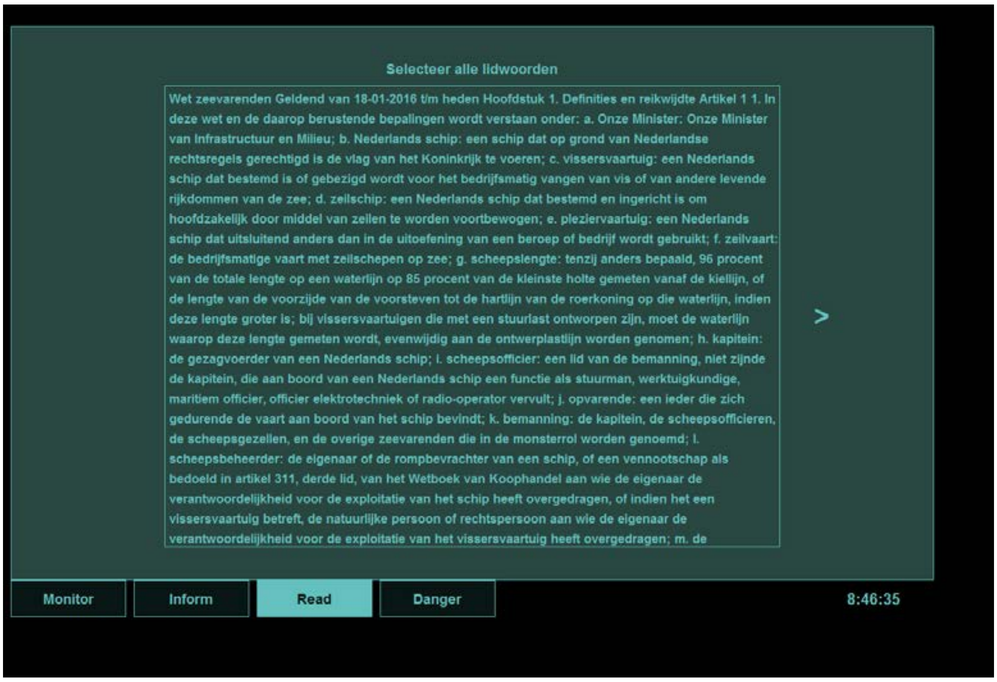


Figure 4-4. Administrator secondary task (self-paced).

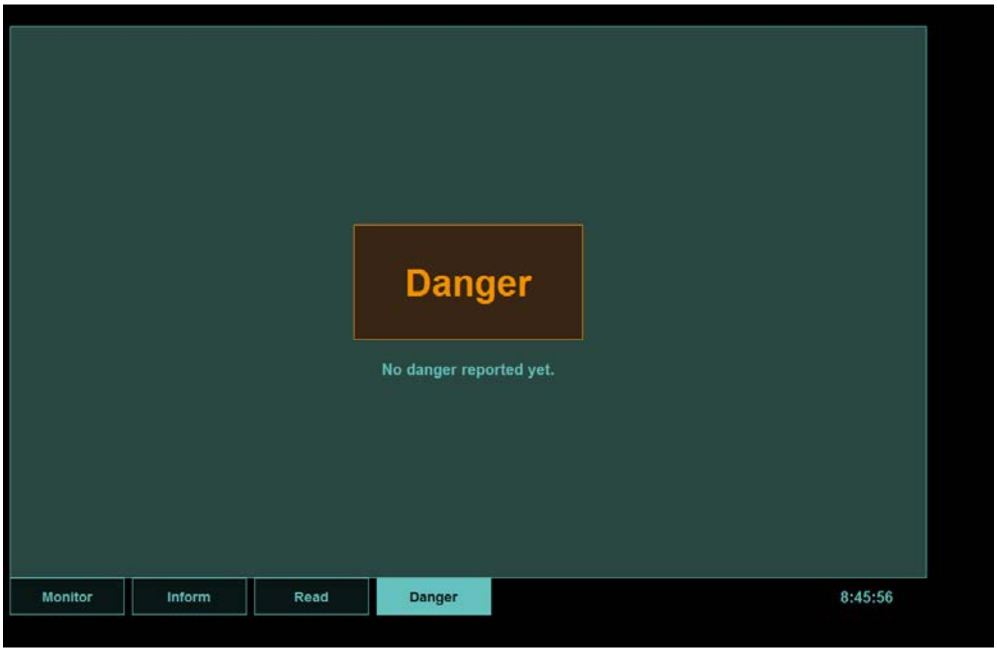


Figure 4-5. Probe secondary task input screen. Hence, ships needed to be detected on a simulated sea environment (see also figure 4-1)

4.3.4 Task instructions

Participants were given general information about the work of a DPO and what variables could have an influence on the state of the ship. Participants had to imagine that they were a DP operator on a ship performing a stationary DP operation, they had to monitor the values for two forms of abnormalities:

Thresholds and high linear in- or decreases of variables (high variation/variability). Participants were then trained pre-task on rules describing when to take action/ which action was best given certain circumstances. Secondary tasks were present: participants had to monitor the outside environment for the incidental occurrence of other ships approaching. There was also a self-paced administrative task: participants had to select certain predefined words in a text. Moreover, participants had to inform management every 60 seconds of the value of a randomly assigned sensor input value. Hence, participants had to divide attention between different tasks: one primary and three secondary tasks.

4.4 Measures

4.4.1 Primary task performance: reaction time

The performance on the primary DP task was determined by the time it took participants in seconds to decide on the best action to take from the onset of reaching a critical threshold for one of the nine variables in the DP simulation.

4.4.2 Primary task performance: division of time between primary and secondary tasks

To see whether support or stress would have an effect on how time was divided between primary and secondary tasks, we measured the number of clicks on each task window and the time spent on each task, and divided this number by the total number of clicks or time respectively.

4.4.3 Trust in support

In general when there is trust in automation, operators would respond to the alarm emerging from automation quickly and without any hesitation. It was measured, therefore, how long it would take participants (in seconds) to return to the primary task, when attention was on one of the secondary tasks, after the emergence of an alarm.

4.4.4 Performance on secondary task: Inform

The performance was determined by the mean time it took participants to input values in the system, measured from the last input. Ideally, this should be 60 seconds. We were also interested in the standard deviation of the interval, whether participants would be able to keep a steady interval during the scenario.

4.4.5 Performance on secondary task: Read

The performance on the read task was determined by the number of correctly selected words. We also measured the total time spent on this task during each scenario.

4.4.6 Performance on secondary task: Probe

We measured the time it took participants to correctly identify an incoming approaching vessel on the simulated outside sea environment, potentially threatening the DP operation. After detection, the incoming ship disappeared.

4.4.7 Situation awareness

To measure situation awareness, the Situation Awareness Rating Technique (SART) was used (Taylor, 1990) (see also, Van der Kleij et al., 2018). The SART is a simple post-trial subjective rating technique.

We chose to use the SART because it is quick and easy to use, requires little training, and, more importantly, is non-intrusive to task performance. Hence, the SART causes no interruption of the natural flow of the task, as opposed to some of the other techniques using simulation "freezes". This was important to us because we did not want to cause any unintended interruptions during task performance.

The SART measures three separate constructs: Attentional Demand (D), Attentional Supply (S) and Attentional Understanding (U). The three component scores were determined with the SART questionnaire (Kennedy & Durbin, 2005), consisting of 9 statements. All items were measured on seven-point Likert scales, in which a score of 1 corresponds to the most negative response to a statement, and a score of 7 corresponds to the most positive response to a statement. Three statements were used to assess demand, 4 for supply, and 2 statements for understanding. For the experiment, the questions were translated into Dutch and adapted to our setting. To calculate overall situation awareness, the mean scores per component were converted with formula (1) (cf. Satuf, Kaszkurewicz, Schirru, & de Campos, 2016):

$$SA = U - (D - S) = U + D - D \qquad (1)$$

With SA being the Situation awareness score, U the summed understanding, D the summed demand and S the summed supply.

4.4.8 Mental effort

To evaluate mental effort the Rating Scale Mental Effort (RSME) was administered once per test session directly after completion of the task (see also Van der Kleij et al., 2018). O'Donnell and Eggemeier (1986) define mental effort as the ratio between the task demands and the capacity of the operator working on the task. Mental workload is high when the difference between task demands and capacity is small. The RSME, originally developed by Zijlstra in 1993, is a one-dimensional scale with ratings between 0 and 150. The scale has nine descriptive indicators along its axis (e.g., 12 corresponds to not effortful, 58 to rather effortful, and 113 to extremely effortful). It is designed to minimize individual differences. We selected the RSME because it is simple to administer, is not intrusive, and at the same time it provides a good indication of the total mental workload (Veltman & Gaillard, 1996).

4.5 Results

4.5.1 Adaptive support manipulation check

Before analysing the data a check was performed to see whether adaptive support was actually invoked during the adaptive support scenarios. We measured this by looking at the number of times critical values were surpassed, whether attention was on the primary task at the time of the critical situation, and, if not, whether support was provided. Indeed, when a critical parameter was surpassed and the operator's focus was not on the primary task, but, for instance, as determined by the eye tracker, was gazing at the outside simulated world to see whether there was an incoming ship, the support was turned on. This occurred on average once every adaptive support scenario. Hence, in the other conditions significantly more support was provided, $F(2,46) = 28.961$, $p = .000$, partial $\eta 2= .557$ (see also figure 4-6).

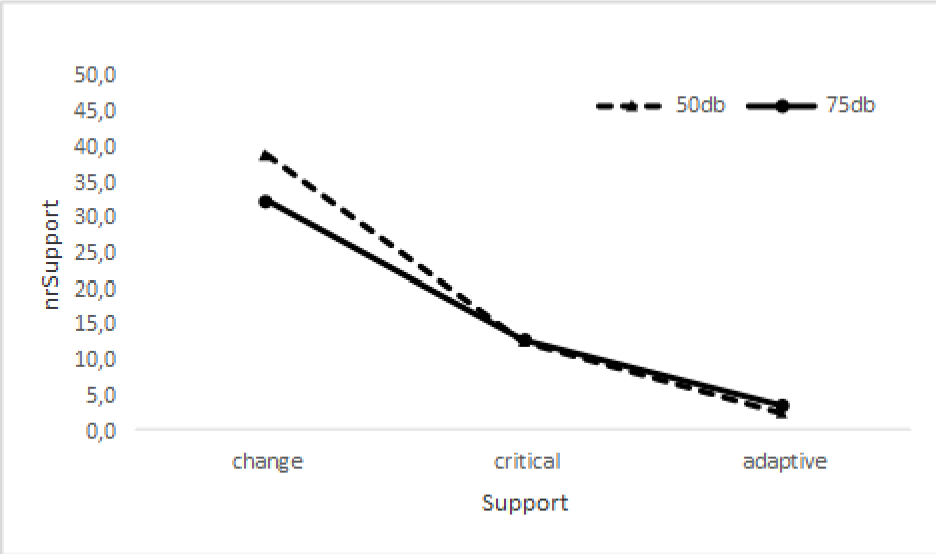


Figure 4-6. Adaptive support manipulation check.

4.5.2 Primary task performance: time to respond

We were unable to perform statistical analysis on this variable: there were too many empty cells. Participants were sometimes unaware of an unfolding emergency situation. Hence, not always were responses logged before the end of the scenario.

4.5.3 Primary task performance: division of time between primary and secondary tasks

A main effect of support was found on the number of clicks on the primary task tab divided by the total number of clicks, $F(3,69) = 7.129$, $p = .000$, partial $\eta^2 = .237$. There was a marginal significant main effect of stress on this variable, $F(1,23) = 3.446$, $p = .076$, partial $\eta^2 = .130$. Further, there was no Support x Stress type interaction, $F(3,69) = 1.544$, $p = .211$, partial $\eta^2 = .063$. Pairwise comparisons shows that support condition 'none' and 'adaptive' differ from 'change' and 'critical', but not from each other. 'Change' and 'critical' differ from each other.

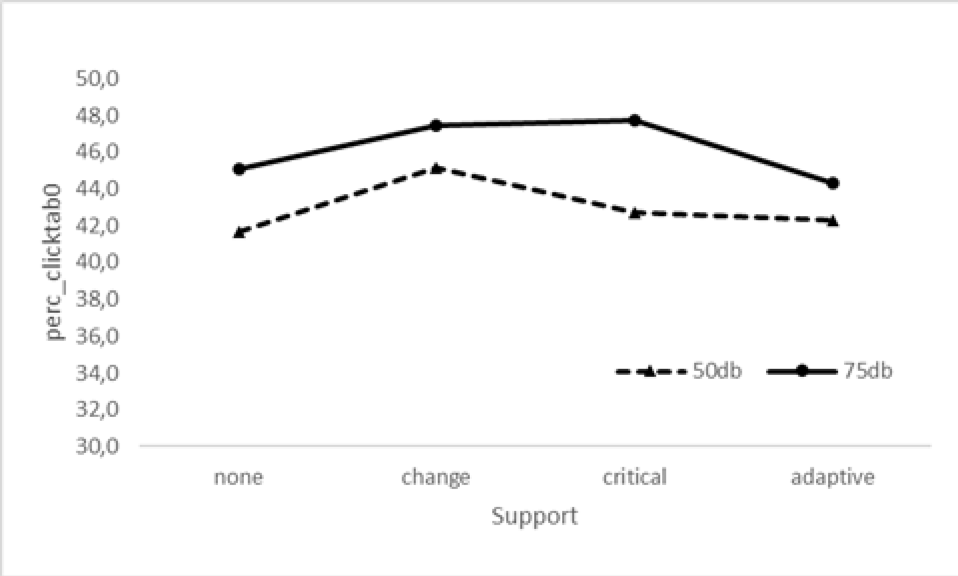


Figure 4-7. the number of clicks on the primary task tab.

Further, a significant main effect of support was found on the total time spent on the primary task tab, $F(3,69) = 3.589$, $p = .018$, partial $\eta 2= .135$. There was, however, no significant main effect of stress on this variable, $F(1,23) = .175$, $p = .680$, partial $\eta 2= .008$. Further, there was no Support x Stress type interaction, $F(3,69) =1.446$, $p = .237$, partial $\eta 2= .059$. Pairwise Comparisons (LSD) shows that support condition 'none' differed from 'change' and 'critical'.

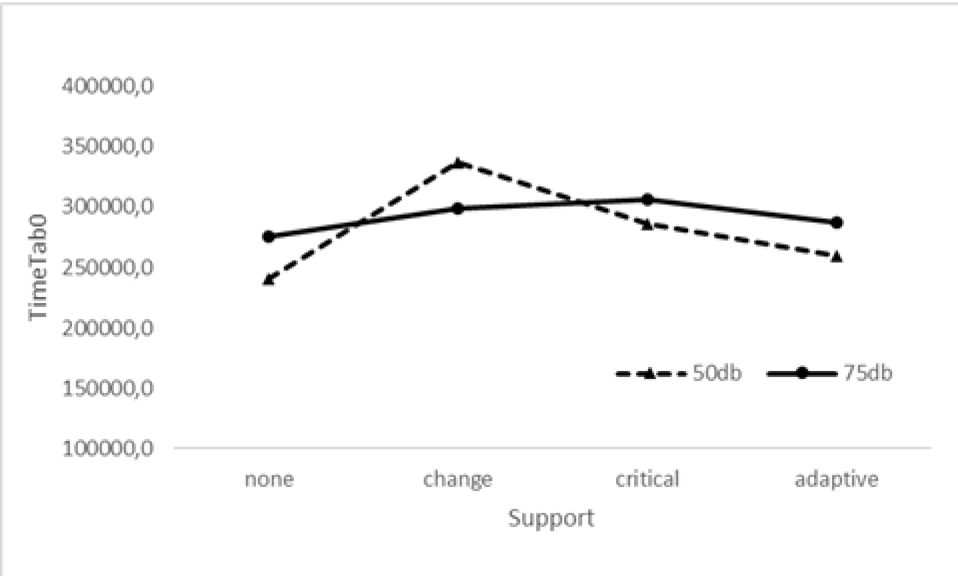


Figure 4-8. Total time spent on the primary task tab.

4.5.4 Trust in support

We were unable to perform statistical analysis on this variable: there were too many empty cells. In the adaptive support condition, not always support was given, resulting in empty cells.

4.5.5 Performance on secondary tasks: Inform

No main effect of support was found on the mean interval time on the inform task, F(3,69) = .725, p = .541, partial η2= .031. There was also no significant main effect of stress on this variable, F(1,23) = 1.062, p = .314, partial η2= .044. Further, there was no Support x Stress type cross-over interaction, F(3,69) = .01, p = .999, partial η2= .00.

A marginally significant main effect of support was found for the standard deviation of the interval time, F(3,69) =2.305, p = .084, partial η2= .091. There was no significant main effect of stress on this variable, F(1,23) = 1.084, p = .309, partial η2= .045.. Further, there was no Support x Stress type cross-over interaction, F(3,69) = .411, p = .745, partial η2= .018.

Pairwise comparisons (LSD) shows that support condition 'none' differed significantly from support condition 'change'(p = .034).

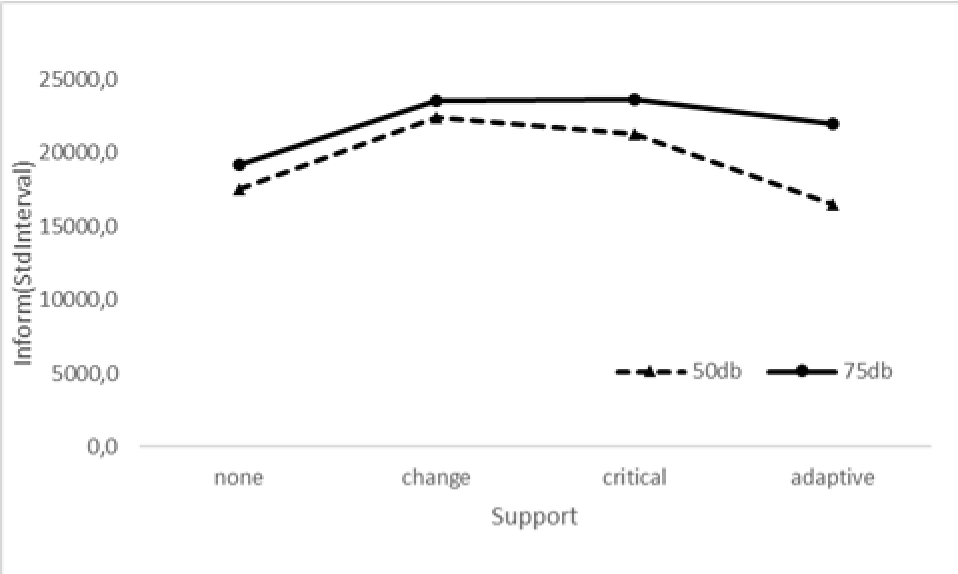


Figure 4-9. The standard deviation of the interval time.

4.5.6 Performance on secondary tasks: Read

A main effect of support was found on number of correctly selected words in the read task, F(3,69) = 4.886, p = .004, partial η2= .175. There was no significant main effect of stress on this variable, F(1,23) = .432, p = .517, partial η2= .018. Further, there was no Support x Stress type cross-over interaction, F(3,69) = .566, p = .639, partial η2= .024. Pairwise comparisons (LSD) show that support conditions 'none' and 'adaptive' differed significantly from 'change' and 'critical'.

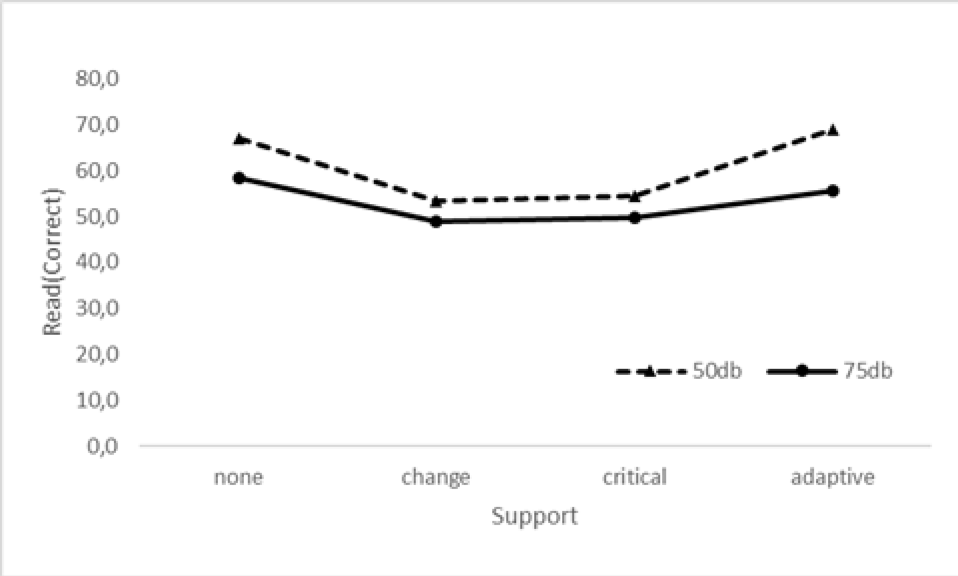


Figure 4-10. Number of correctly selected words in the read task.

4.5.7 Performance on secondary task: Probe

No main effect of support was found on the mean response time for signalling a probe (incoming ship), F(3,60) = .281, p = .839, partial η2= .014. There was also no significant main effect of stress on this variable, F(1,20) = 1.402, p = .250, partial η2= .066. Further, there was no Support x Stress type cross-over interaction, F(3,60) =1.089, p = .361, partial η2= .052.

4.5.8 Situation awareness

No main effect of support was found on SA, F(3,69) = .826, p = .484, partial η2= .035. There was also no significant main effect of stress on SA, F(1,23) = .007, p = .934, partial η2= .000. Further, there was no Support x Stress type cross-over interaction, F(3,69) = .190, p = .903, partial η2= .008.

4.5.9 Mental effort

No main effect of support was found on RSME, F(3,69) = 1.887, p = .140, partial η2= .076. There was also no significant main effect of stress on RSME, F(1,23) = .011, p = .919, partial η2= .000. Further, there was no Support x Stress type cross-over interaction, F(3,69) = .373, p = .727, partial η2= .016. It should be noted, however, that the overall score was quite high (M=64.20; SD=25.66).

4.6 Discussion

This study was designed to investigate how adaptive automation might be used to address operator underload states in highly reliable automated systems. Our first goal was to design an adaptive algorithm that assesses whether the operator is paying attention (or not) and decides, based on the situation and, hence, the need for operator involvement, whether the operator needs to be alerted (or not).
Our data revealed that we were successful, on the basis of operator's eye movements, in resolving whether the operator was paying attention to the primary task, and, consequently, providing support when this was deemed necessary.

We were less successful, however, in proving the benefits of adaptive automation on overall system performance. Due to an unforeseen technical problem with logging responses, which was not present during pilot testing but was introduced with an experiment software update, we were unable to test for effects of support on operator response time, our primary performance measure.

Other measures showed no differences between the no-support condition and the hybrid adaptive-support condition. No effects whatsoever were found of support on mental effort or situation awareness. Our study did confirm, however, the detrimental effects of ill-designed support on performance: The static and the critical event support conditions, both causing an abundance of alarms, most of them false, decreased overall system performance (see also van der Kleij et al., 2018). So, the conclusion could be that hybrid adaptive support prevented the operator to become overloaded with alarms, thereby allowing him or her to maintain stable system performance on the relatively short duration of the task.

An interesting finding is the absence of stress effects on performance. This may seem surprising at first glance since noise impairments are common (Matthews et al. 2000). Since no negative effects were observed on critical outcome measures, this may suggest that operators adapted to the stressor by using effective compensatory strategies (Hockey 1997). According to Hockey's theory, compensatory strategies are typically found to increase effort expenditure, diverting attention from secondary to primary tasks, or increasing risk taking. However, post-hoc analyses revealed no proof of participants successfully adapting coping mechanisms during high stress scenarios. No differences were found between stress conditions in the time that was allotted by participants to the primary task as compared to time devoted to secondary tasks. A possible explanation is the complexity of the task. Some participants complained that the task was too complex and difficult to perform well. Indeed, we observed that participants were performing at low levels, perhaps leaving no room for further decrements in performance as a consequence of stress.

5 Operator Cognitive State modelling using unobtrusive measures

To offer adaptive automated support at the right time it is important for the automation to be aware of the cognitive state of the human operator. This is to prevent the phenomenon of ‘disruptive automation’, where support only adds to the already high level of workload and further impedes the operator’s ability to respond correctly. The experiment described here was intended to make advances in operator state modelling. The goal was to determine the operator’s level of Situation Awareness based on his problem-solving strategy. This was done by comparing the operator’s course of action to the optimal solution as determined by an expert operator.

This was done in a simulated Dynamic Positioning (DP) environment. Here, the DP operator and his bridge crew were confronted with several failure modes during several training exercises. They were graded by an instructor based on how well they handled the failures.

An experiment was created around the DP training in order to model the operator state.

This was done by facing three challenges; measuring the operator, determining the optimal solution, and comparing operator actions to the optimal solution.

5.1 Method

5.1.1 Measuring the operator

The first step towards drawing conclusions regarding operator SA was to enable the accurate monitoring of operator behaviour.

The measurements that were taken for this experiment always focussed on the person that was assigned as Dynamic Positioning Operator (DPO), other bridge crew were not specifically monitored, unless they interacted directly with the DPO.

To get a good overview it was important to record all the system interactions and the communications relevant to the DP task. This included the direct interactions of the DPO with the DP-system, including the information that was visible on the screens at any given time, what the operator was looking at, and also which settings were changed on the DP system. Secondly the DPO had interaction with his bridge crew as they discussed relevant information and possible courses of action. Finally there was also communication between the bridge and third parties located outside the bridge (e.g. machine room, platform) that could provide vital information to the DPO.

Several sources of information were recorded simultaneously. Each of the information sources will be described separately, however, it should be noted that great care was given to ensure that all the information sources were time-synchronised and any phase errors were prevented.

5.1.2 Video/audio

The DP desk consisted of a master and a slave screen. The same information was available to both screens, however the operator had to make a selection as to which information was actually visible. Each screen was divided into two halves which could be altered by the operator to display any information they preferred, essentially giving 4 changeable areas of information. For the analysis it was important to know which information was available to the operator at any given time, therefore the DP screens were recorded as a video stream.

To achieve this we used a video-logging system that duplicated and recorded the image of both DP desks at 60Hz at a slightly reduced image quality (analogue) of around 480p (DVD quality). One of the video streams also included audio.

The audio was recorded using a microphone situated on top of the DP desk, i.e. central on the bridge. This allowed the recording of DPO contemplation, but also bridge interaction with the DPO and radio communication with third parties outside of the bridge.

5.1.3 Eye tracking

The video stream of the DP system recorded the information that was displayed at any given time. To measure where on the screen the operator was looking an eye-tracking system was used.

For this experiment the SmartEye eye-tracking system was chosen, using a 4 0.3MP camera setup. To achieve the highest possible gaze resolution of the main desk a combination of both 6 mm and 8 mm lenses was used to enable a wide field of view for the eye tracker to allow tracking while the operator moved naturally behind the DP desk, but also a high gaze resolution when the operator was standing at the default working position, i.e. directly in front of the master DP desk. Figure 5-1, displays the approximate viewing angles and range of the eye-tracking system. The areas marked in orange are where the operator could stand while being tracked. Figure 5-2, provides a picture of the actual setup that was used.

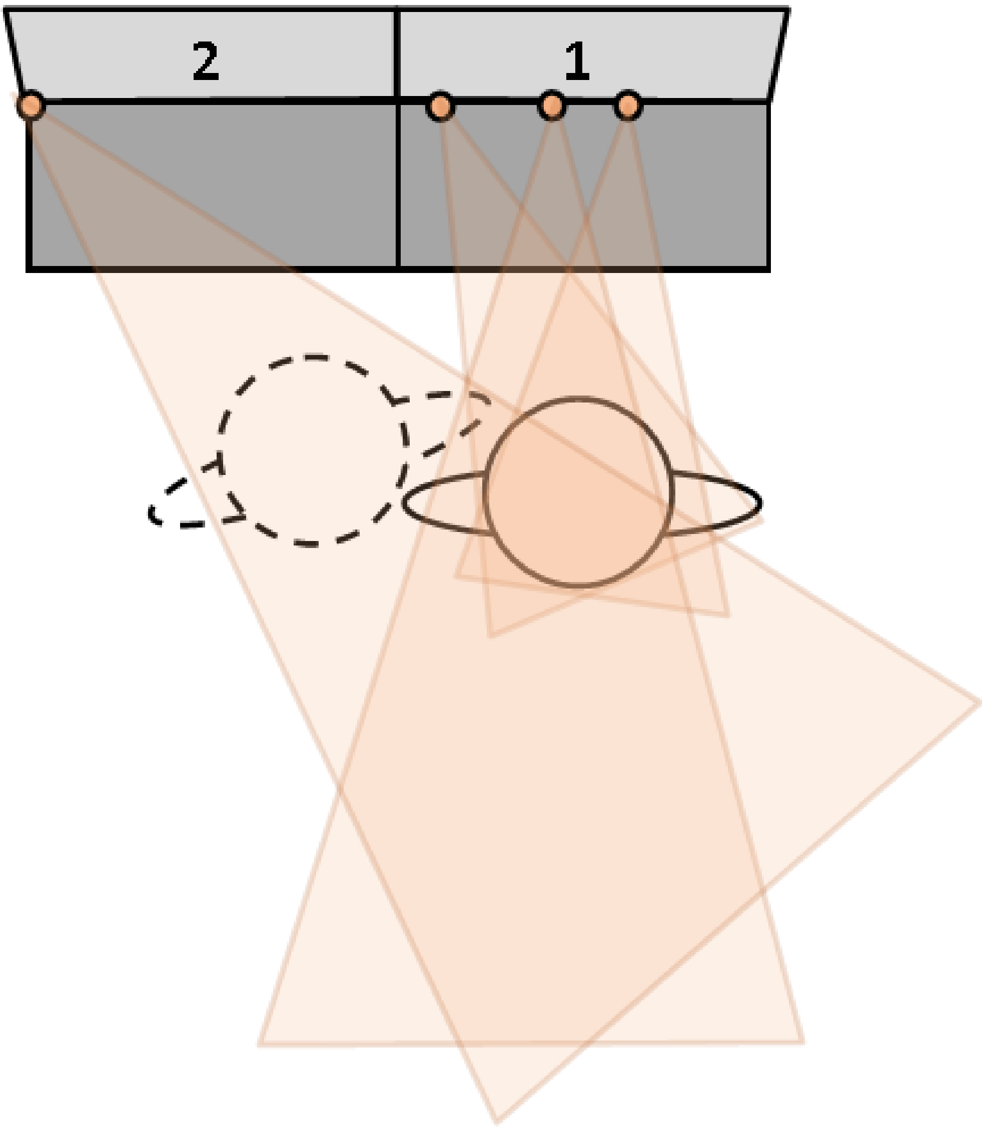


Figure 5-1. Approximate model of camera setup design, allowing for some freedom of movement.

Figure 5-2. Picture of setup in simulator. In the middle the eye-tracking cameras.

The eye-tracking equipment had to be calibrated to each individual user to achieve the highest possible accuracy. As it was not allowed to interfere during the execution of the training scenario the equipment was calibrated to each individual beforehand, the calibration was then saved and could be switched when the trainees switched roles on the bridge. Switching calibrations was controlled by the experiment leader.

To enhance the richness of the eye-tracking data a 3D model was constructed within the SmartEye software representing the layout of the DP console, including both screens and both control panels (Figure 5-3). In addition to coordinates, this allowed the system to distinguish gazes to specific areas of interest.

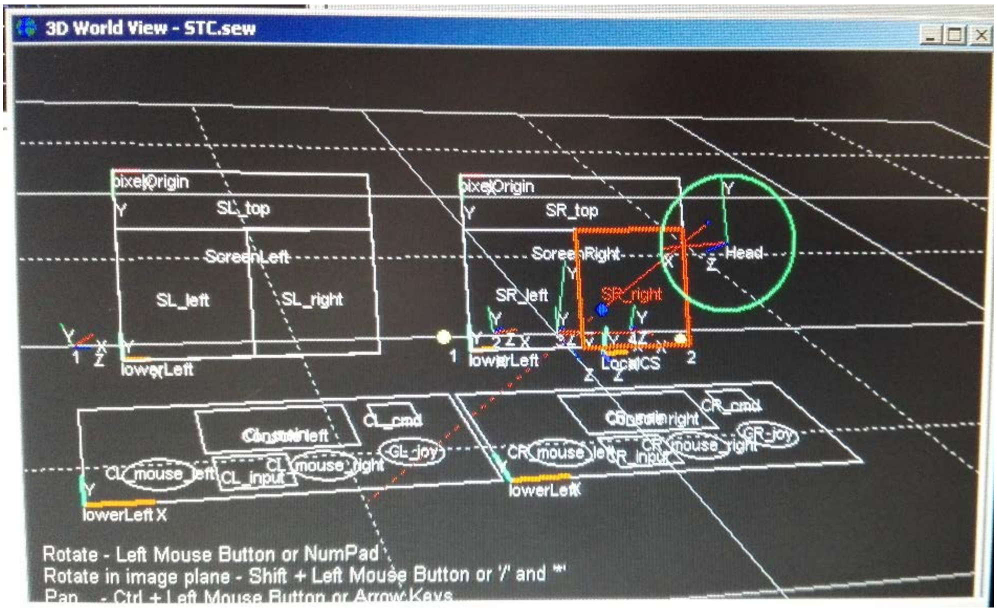


Figure 5-3. Image of 3D-model including head position (green) and gaze direction (red).

5.1.4 Key logging

Even though the DP console provides a lot of on-screen feedback of buttons pressed and settings chosen, a direct measure of key-presses was desired. Due to the custom architecture of the DP console it was not possible to tap the signal and record anything meaningful. Therefore it was opted to use a wide-angle video camera to record keypresses. This was recorded in sync with the video feed of the DP console for easy analysis.

5.1.5 Reconstructing the operator solution path

A Cognitive Work Analysis (CWA) procedure was used to reconstruct the actions taken by the operators towards resolving the encountered failure modes. CWA means that at a cognitive level one looks at information elements which should trigger certain actions (if-then relationships) or to put it differently, which information justifies an intervention.

As input for the CWA the recordings were used that were made during the simulation. This was enriched by using the observations from the experiment leader during the simulation, as well as the instructor's commentary, and the debriefing following the training session. This resulted in a step-by-step account of the actions that were taken by the DP operator for every failure mode that was presented during the simulation. The result of this effort will be discussed in section 5.5.

5.2 Creating the optimal solution-path

The optimal solution path was created prior to the training in collaboration with the DP instructor who also conducted the training. All the failure modes that would be presented during the training were known to the experiment leader in advance. A CWA was made for all of these failure modes. The results of these efforts can be found in Appendix B.

5.3 Systematic comparison of operator actions to optimal solution

After creating an account of the optimal path, and reconstructing the actions of the operator, the two were put side-by-side to see where the operators deviated from the optimal solution path. The result was put in a graphical representation (figure 5-4) that will be discussed in detail in the results section.

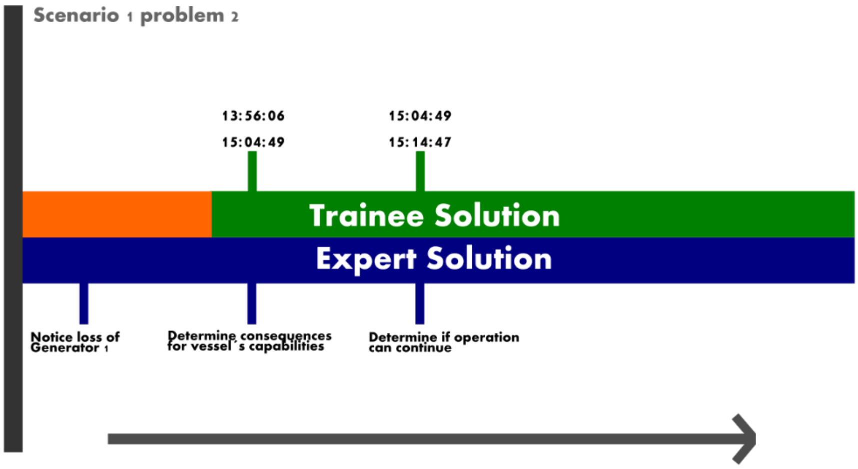


Figure 5-4. Example of trainee solution path deviating from expert solution

5.4 Experimental setup

5.4.1 Location

The experiment took place at the 'Scheepvaart en Transport College' (Shipping and Transport College) in Rotterdam, the Netherlands. The STC owns a high-fidelity ship simulator that is frequently used for education, training and certification purposes for individuals ranging from students to highly experienced professionals.

Measurements were taken during a training course for Dynamic Positioning Operators. This course is given to experienced seamen who are getting additional certification to use a Dynamic Positioning system during commercial operations.

5.4.2 About the training

The DP advanced course took a total of four days to complete. Each day followed roughly the same schedule.

First, the participants were given time to prepare for the simulated mission. They were given an assignment and a map of the situation including the position of their own vessel and any obstacles. The goal during preparation was to devise the optimal way to complete the mission, taking into consideration (threats to) safety, efficiency, and which sensors / actuators to use throughout the process.

After the preparation phase the trainees moved to the simulator where the trainer assigned roles to each of them. All trainees were on the bridge simultaneously and assisted each other and could ask each other questions, but only one person was assigned as DP operator at any one time and was in control of the executive decisions.

Throughout each day of training the aim was to assign every trainee each role once. Therefore on average there were between three and four different DP operators standing behind the desk.

During the execution of the operation the trainees were confronted with a range of failure modes. These were meant to test the capability of the trainees to deal with sudden unexpected situations, e.g. thruster failure, loss of position reference system. The observations that were done for this research focussed on these incidents. Of special interest was the way in which the assigned DP operator noticed the failures, identified them, and came to a solution.

After the mission was completed the trainees were debriefed and errors (if any) were discussed with the trainer. All phases of the training, including debriefing, were attended for research purposes.

Out of the total of four training days, measurements were taken during the first three days. The final day was reserved for examination, to avoid any interference or adverse effects of monitoring on the successful completion of the training by all trainees this day was excluded from the research.

The following paragraphs describe each of the three training days that were attended (training descriptions and failure modes by courtesy of: STC B.V., The Netherlands).

5.4.3 Day one

5.4.3.1 Mission Description

"*Transport a generator set from Platform A SE side to Platform C. Weight of the generator is 1 tonne, which means that the platform cranes can handle this generator at their maximum outreach. Pipe lay stingers are not mounted, so the aft deck is clear.*"

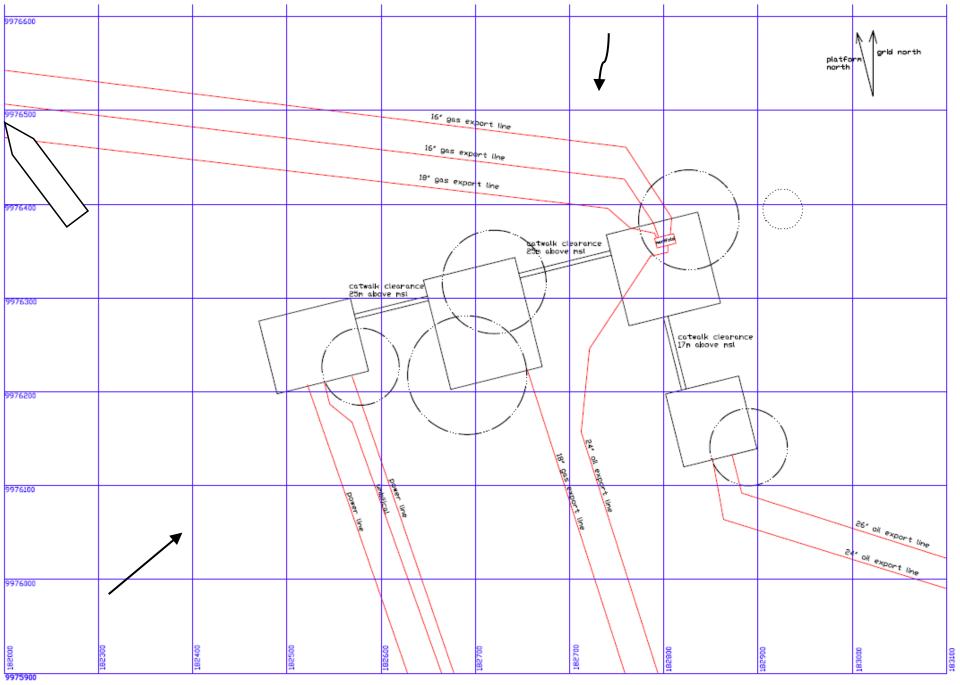


Figure 5-5. Sea chart describing position of ship and platforms.

5.4.3.2 Failure mode description

1. During approach to platform A gyro 1 starts drifting (see alarm message) DPO should check on sensor page and inform survey or electrician to check the gyro. System is automatically switching over to gyro 2.
2. When underway to C platform generator 1 fails DP operator should check power screen and check capability plot. Before starting 2nd approach. Decision to make continue or stop operation.
3. During approach C platform wind sensor erratic due to turbulence around platform sensor must be deselected by DP operator on sensor page.

5.4.4 Day two

5.4.4.1 Mission description

"*Underneath Platform C is a subsea manifold. From this manifold, three pipelines go in a north westerly direction. A fourth pipeline goes in a south westerly direction. This pipeline has to be inspected by divers, as there is a suspicion of leakage. Very small amounts of oil appear on the surface underneath the platform every now and then.*

This means that the divers only need to inspect the segment of the line that lies underneath Platform C. The diving superintendent has been very clear that he wants the Challenger on the west side of platform C, as close to the platform as possible."

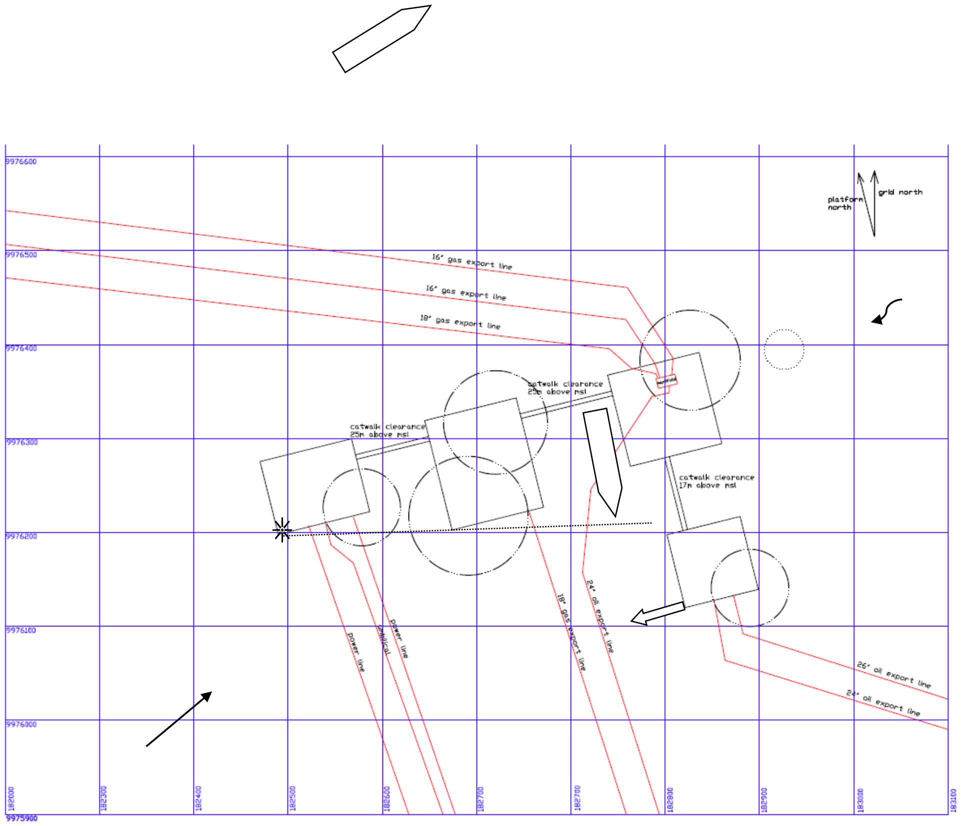


Figure 5-6. Sea chart describing position of ship and platforms.

5.4.4.2 Failure mode description

1. Sail south from starting position, and outside of 500m zone fill in DP checklist and request permission to enter 500m zone, and ask platform to switch on required position reference systems. (artemis,radius)
2. Start approach, during approach DGPS1,DGPS2 and Artemis on SW A platform available. This stops when no longer visible as soon as antenna disappears behind B platform (see dotted line). Because 3 reference systems have to be online it has to be known why artemis dropped by looking at the reference systems page on the DP console. Next, a third system will have to be used. The choice can be made between a HPR transponder on the seabed or a Radius transponder on the platform. This is dependent on the heading of the ship.
3. During the approach Bow Tunnel 2 will fail. To remain redundant the Bow Azimuth will have to be switched to bus 2, and the power screen has to be selected to be able to see the settings.
4. When a choice is made and the ship is at its final location, the diving operation can commence. It is a bell-dive with three divers, of which 2 will perform an inspection, and the third will remain in the bell.
5. When the dive is in progress, DGPS1 and 2 start drifting. This can be noticed by the pink bars behind the reference systems, but is better visible on the reference systems page. Reduced weight should be applied to DGPS to reduce the impact of the issue.
6. Engineering reports that bow tunnel 2 is available again and ready to be switched on. If the answer is yes the bow tunnel will provide full power as a result of a faulty repair.

The ship will be in trouble if the emergency stop is not activated on time. When it is deselected within the DP system it is considered a wrong action.

5.4.5 Day three

5.4.5.1 Mission description

"*Perform a cable lay operation, starting at platform . This is the plan:*

1. Perform a position check on the easterly calibration transponder.
2. Come underneath the crane of platform E with your stern deck. You will receive a pre-installed cable from platform E. This is already properly connected to platform E.
3. Your deck-crew will make the connection between the platform end of the cable and the cable on your reel. Once this is finished, the cable will go overboard from the stern cable lay position.
4. While the deck-crew is making the connection, you can re-position the vessel to the starting point, from where the cable has to be laid: N67.93537 E40.6267.
5. Please note that this position is outside the reach of the platform crane. Ensure that the platform crane is clear from your deck before you start to move and keep close communication with the deck foreman to avoid extreme tension on the cable.
6. Prepare the route in the route editor the cable should be 20m south of the control line CL_202 and deployed with HPR B23 transponder on the end on template A for pick up at a later stage of the project. cur
7. The cable should be laid in a +/- 5m corridor."

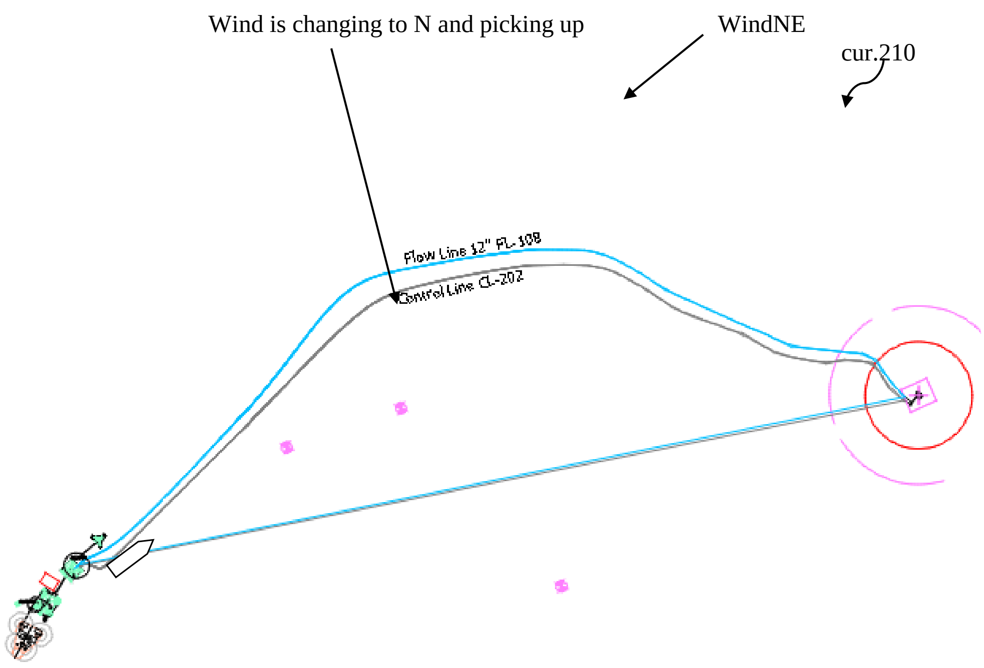


Figure 5-7. Cable laying trajectory.

5.4.5.2 Failure mode description

1. During approach to platform E, bow azimuth fails cannot be repaired. Capability plot should be checked
2. During laying wind is changing and picking until vessel cannot hold heading, Cable lay must be stopped and heading change carried out crab angle maximum +/- 30 degrees.
3. Decision must be made if ROV must be recovered before weather decreases vessel must create leeside before recovering ROV, heading change required. (minimum roll)

5.4.6 Participants

The advanced Dynamic Positioning training was given to four people who had all completed a basic DP training and 60 days of active DP working experience at sea. All participants were male and were not especially selected for the experiment but were participating in the DP advanced training that was subject of the observation.

5.4.7 Procedure

An experiment day commenced with the operators receiving their mission briefing from the instructor. Following this brief, the bridge crew made a planning regarding the steps to take in achieving the mission. In addition they reviewed which reference systems to use along the way, and when to communicate with whom about what. It was not allowed to be present during this phase to prevent disruption of the training process.

After the mission planning they moved to the bridge simulator where the experimental setup was located. Before the simulation started the eye-tracking equipment was calibrated to each user individually. As it was not allowed to interfere during the training session, each user got a saved profile on the eye tracker computer, which was switched by the experiment leader as the roles on the bride were reassigned. This is due to a limitation of the eye tracker which can only accurately track one person at a time.

The training scenario was then initiated by the DP instructor and the experiment leader was seated next to the instructor in a separate room. During the unfolding of the scenario notes were taken regarding important events in the scenario, and key decisions made by the bridge crew and of course the DPO.

After the training, the operators were debriefed by the instructor. This debrief was witnessed by the experiment leader and provided additional insight into the reasoning behind certain decisions that were made.

5.5 Results

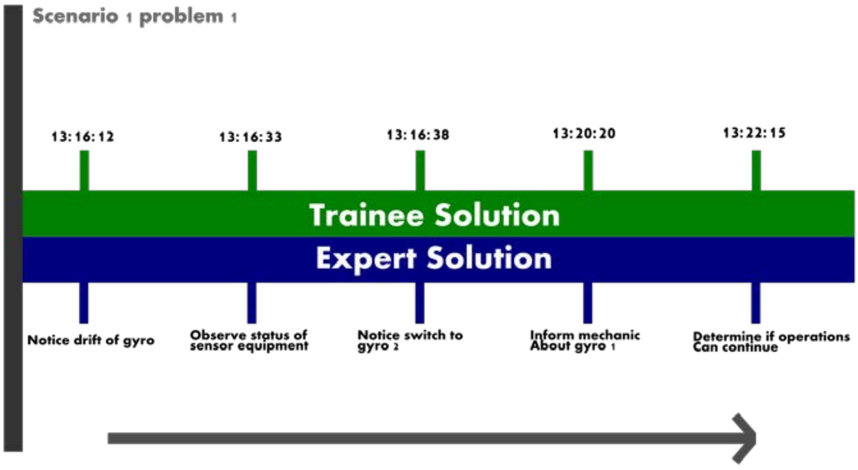
5.5.1 Comparison of optimal solution to operator actions

Results are displayed for each training day and failure mode separately. A brief recap of each mission and failure mode is given here, they are described in more detail in paragraph 5.4.2, and appendix B, where the optimal solution path is given. In this analysis a visual comparison is given between the optimal solution path and the chosen actions by the operators. The optimal solution path is marked in blue, including a listing of the correct actions to be taken. The solution path of the operators is marked in green when they followed the same solution path as the expert. A temporary deviation from the optimal path is marked in orange, and a complete deviation from, or omission of, the correct responses is marked in red (Houtkoop, 2017).

5.5.2 Day 1

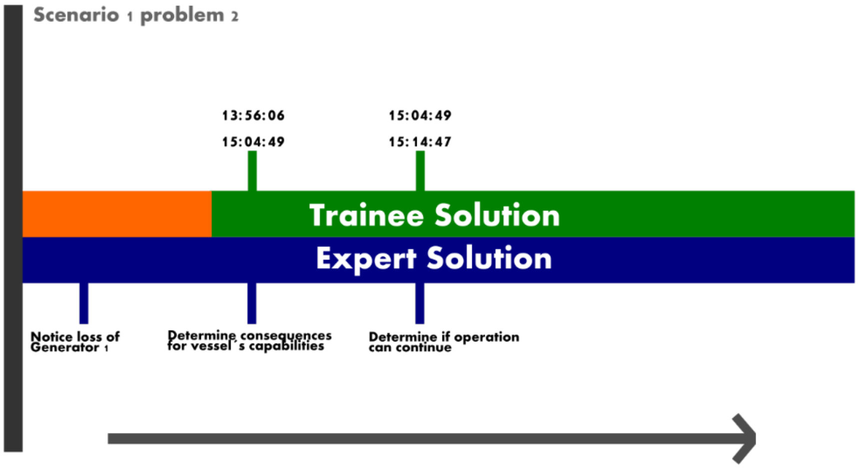
This mission involved transporting a generator from one platform to another. During the exercise three failure modes occurred.

5.5.2.1 Failure mode 1 – Drift of gyro 1



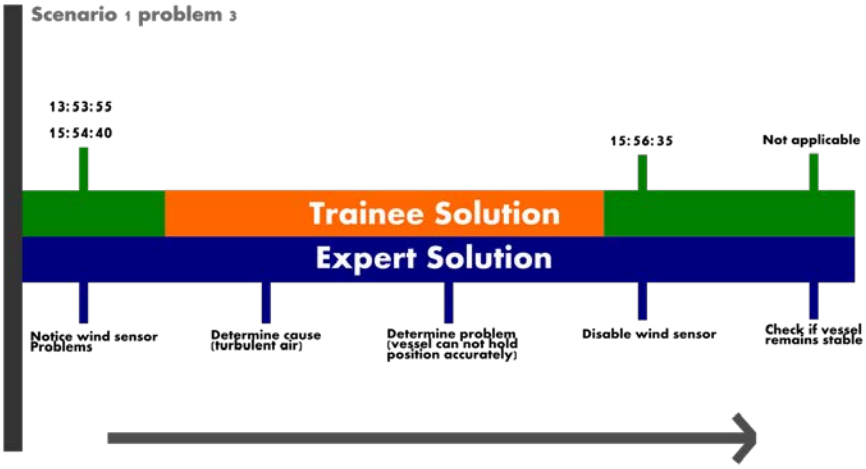
Problem assessment:
Trainees reacted to the faulty gyro, the problem was noticed and actions were taken to eliminate the problem. The vessel was not at risk during this problem.

5.5.2.2 Failure mode 2 – generator failure



Problem assessment:
The problem with the generator was not detected immediately by the trainees. They reacted to the generator failure only after a second generator failed as well. The actions that followed after noticing the failures were correct and in accordance with the expert solution.

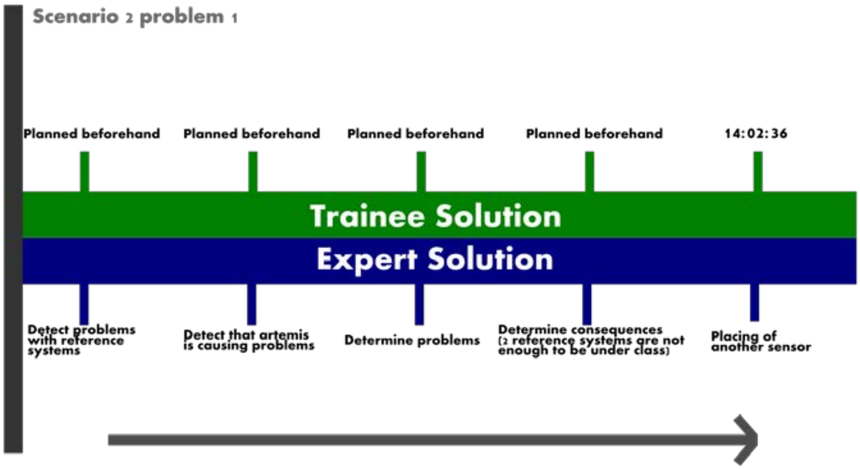
5.5.2.3 Failure mode 3 – turbulent air disrupting wind sensor



Problem assessment:
The trainees noticed the wind sensor problem, but did not notice that the problem was caused by turbulent air circulating around the platform, which is a common occurrence. Eventually they disabled the wind sensor because the error kept returning. However, by the time they disabled the sensor the vessel could already leave the turbulent wake of the platform and the problem would have dissipated.

5.5.3 Day 2

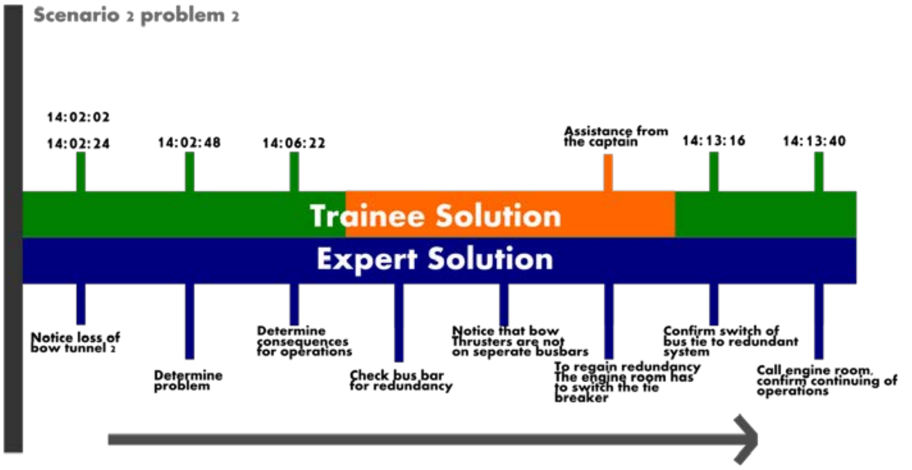
5.5.3.1 Failure mode 1 – Artemis obstructed by platform closeby



Problem assessment:
During mission preparation the trainees already noticed that the Artemis transmitter would become obstructed during the approach. Therefore the problem was not a surprise to them, they reacted beforehand to the problem and dealt with the problem by initiating the RADius transponder.

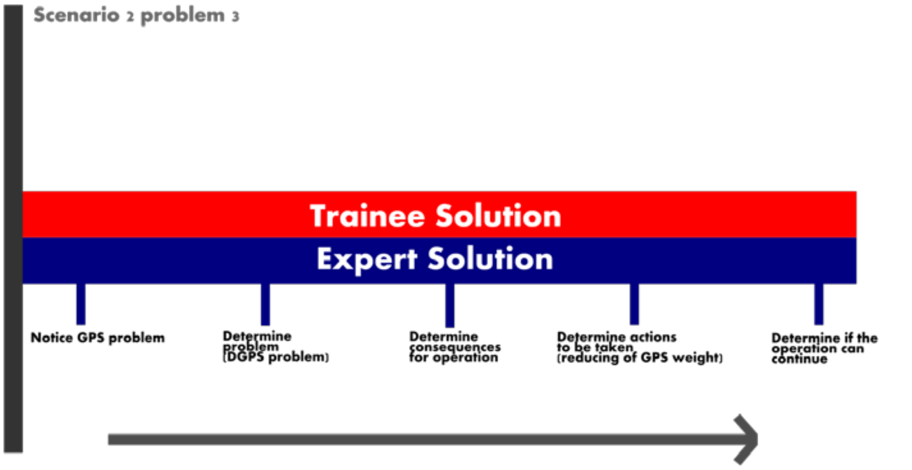
Footnote: the Artemis system uses microwaves to determine the absolute distance and relative angle between two artemis stations.

5.5.3.2 Failure mode 2 – bow thruster failure



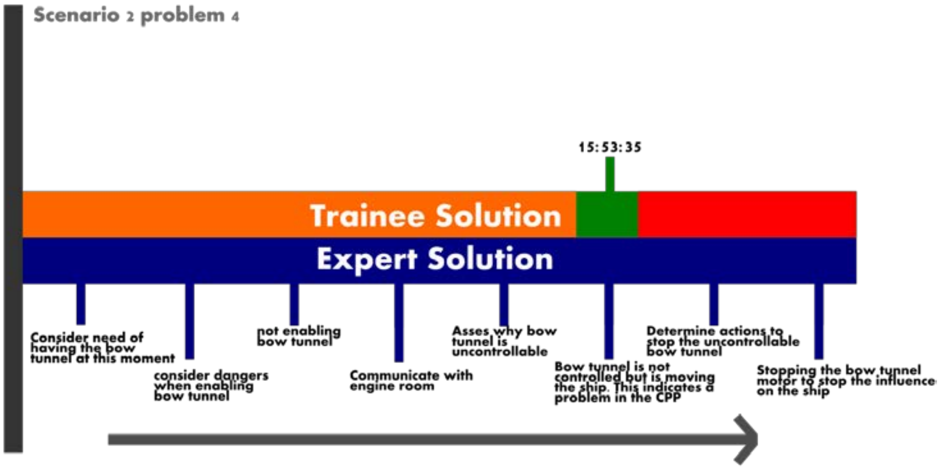
Problem assessment:
The trainees reacted to the bow tunnel problem, afterwards they were unsure about the consequences for the class of the vessel. The instructor had to step in during this scenario problem, the instructor fulfilled the role of captain and discussed being under class with the operators. They were unaware of the switching of the tie breaker to regain redundancy, after explaining this the operators continued.

5.5.3.3 Failure mode 3 – GPS drift



Problem assessment:
The trainees noticed a problem with the reference systems only after the system already rejected the RADius transponder. They did not recognize this as a problem with the DGPS drift. The result was that the operators acted according to a RADius error and not a DGPS error.

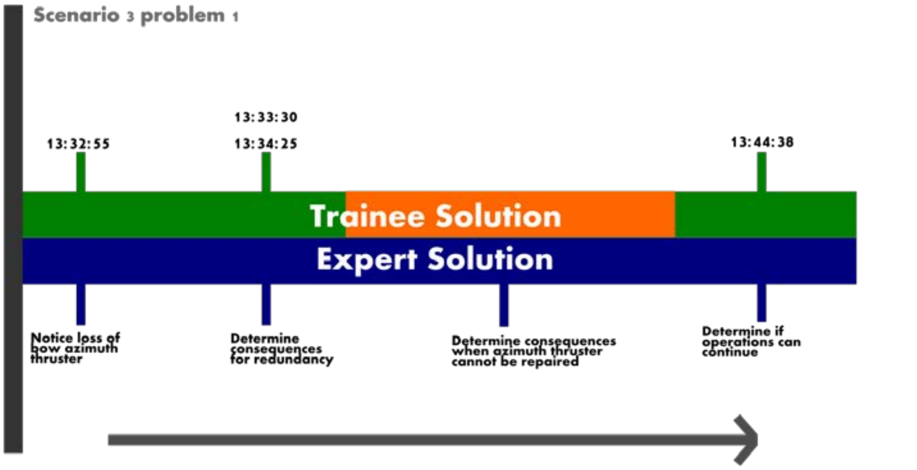
5.5.3.4 Failure mode 4 – Re-engaging bow thruster after repair



Problem assessment:
During this problem situations there were three very clear errors: The bow thruster was re-enabled during a high risk phase of the operation. The other mistake was that starting bow tunnel 2, happens without corresponding with the DP officer. Both the DP officer and the navigational officer raise concern. This is discussed with the radio officer, however the bow tunnel is already being started.
Due to a faulty repair the engine provided full thrust in one direction, pushing the vessel out of position. This might have been mitigated by engaging the emergency stop procedures. However none of the operators recollected this option which eventually led to the vessel being uncontrollable and colliding with the platform.

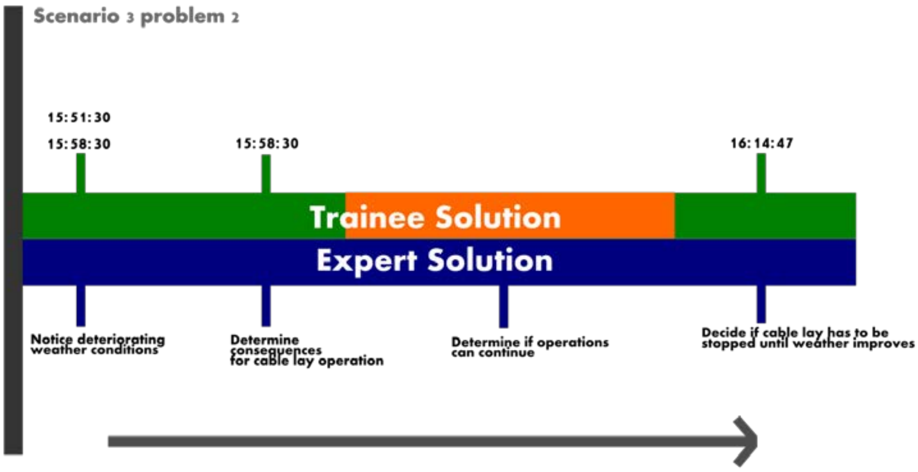
5.5.4 Day 3

5.5.4.1 Failure mode 1 – Azimuth thruster failure



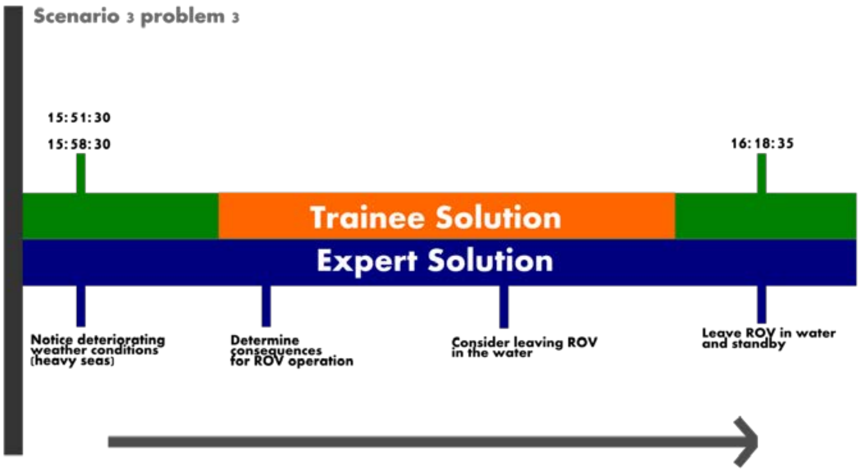
Problem assessment:
The trainees noticed the problem with the azimuth thruster and reacted to it, the failure caused no additional problems and the operation could continue. The difference with the expert solution was the making of a capability plot, the trainees did not do this and therefore did not know the negative effects on the vessels performance.

5.5.4.2 Failure mode 2 – sudden change in weather



Problem assessment:
The trainees conducted the cable lay until the weather deteriorated, instead of considering alternative options they decided to stop operations and wait for the weather to clear up.

5.5.4.3 Failure mode 3 – recovering ROV due to weather



Problem assessment:
The weather deteriorated during the cable-lay operations, the operators eventually stopped the cable lay when the weather conditions were too bad to continue. The ROV was left in the water on standby and was not recovered, when the operation was stopped recovery was no longer possible due to the wave height. ROV crew is not informed during operation, only informed afterwards when the choice for standby has been made. By leaving the ROV in the water the operators risk the potential loss of the ROV.

5.6 Discussion

5.6.1 Operator state modelling
One of the challenges of this research was to devise a way to unobtrusively measure the cognitive state of the DP operator. By integrating various methods it was possible to determine during post-analysis what the operators were doing at any given time. This allowed for the distinction to be made between whether an operator simply did not notice the issue (low saliency) or whether he was actively engaging in problem-solving strategies. It was also possible, by monitoring the operator gaze through use of the eye tracker, to determine what information the operator was searching for while attempting to solve the issue. This allowed for a solution path to be reconstructed and compared to that of an expert operator. The solution path gave insight into when and how a chain of errors started. This ability will enable developers of future support systems to analyse when and how support should be offered, i.e. during detection, or assessment of the problem.

5.6.2 Sources of errors

From the performance of the operators on the various failure modes it became clear that several different causes of errors could be distinguished. While the precise nature of the errors varied, similarities can be found among them.
For example multiple errors could have been prevented if the DP system had provided more salient alarms. The operators missed the generator failure because they did not notice the drop in power availability, even though it was visible on their screen. Alarm saliency can be improved through the design of the DP interface.
For the purpose of this study it is of more interest to consider the errors that were made for reasons related to the human operator, e.g. cognitive underload, overload, distraction, lack of Situation Awareness (SA).
Several errors were made as a result of a lack of contextual awareness.
For example, when the DGPS drift occurred, none of the operators considered that even though their system seemed to show that RADIUS sensor to be drifting, it was far more likely that the DGPS values were drifting as a result of their proximity to a platform. As a result they spent a lot of time trying to fix a system that was functioning perfectly, and left the actual problem unattended. From these observations it cannot be concluded whether this was due a knowledge gap, or whether this was the result of perhaps cognitive overload while attempting to solve the imminent crisis. For the error that was made by omitting to engage the emergency stop for the faulty bow thruster it was more evident that this was a result of cognitive overload. The operators had practiced the correct procedure only hours prior to the incident occurring, however they were unable to recollect the correct course of action when the engine failed. At the time that the error occurred the situation was very stable and operator workload was low. When the engine suddenly malfunctioned the operators were required to quickly restore their level of SA to analyse the issue and to come up with a solution. During the simulation this process took a long time as the operators were likely suffering from Out Of The Loop (OOTL) related performance issues. In fact, the correct solution was not found until it was already too late.

5.6.3 Suggestions

While some errors can be prevented by simply improving the design of the DP interface, most errors that were made were not the result of failures going undetected, but the result of the operators being unable to come up with the correct courses of action to mitigate the situations. These types of errors provide an opportunity for future systems to support operators when their tasks are highly demanding. The techniques described in this paper can serve as a basis to refine operator state modelling and to start moving from post-analysis to (near) real-time analysis of the operator state. This capability will allow a support system to provide adaptive operator support depending on the context of the situation and the state of the operator.

These findings underline the importance of human aware computing and human aware AI. The fact that technical systems can share their SA with humans plays a significant role in operator support in unexpected situation. Hence, it also works the other way around. Because it works in both directions, it facilitates shared SA.

These findings also, illustrate that it is potential beneficial to extent the functionality of the IOSS to decision making support. Hence, to provide support in situations in which operators are being unable to come up with the correct courses of action to mitigate the situation. Decision support does not mean that IOSS should come up with the ‘best option’ or ‘the answer’ (if it was that straightforward the reaction could be automated as at it-then rule). Support could consist in helping to reduce the problem space by excluding systems aspects as source of failure, or to help with generating alternative courses of action to mitigate the situation, or with providing help with the question whether actions will be effective in solving the situation given the manoeuvrability envelope.

6 Overall discussion

On the basis of the literature study that formed the basis for the study on adaptive maritime automation, the conclusion can be drawn that a dichotomous way of division of tasks between humans and autonomous systems is not satisfactory, leads to inefficiency and impairs the ability of operators to take mitigation measures in case of operational challenges. Furthermore, as the automation becomes more reliable and more robust, the dichotomy problem will become even greater because the chance that the human operator is aware of critical information under these circumstances and can take over the control manually when necessary, will become smaller.

The approach that has been followed in this project can be characterized as an experiment to ‘replace’ the current dichotomous way of working with a joint human-automation framework. Within a joint human-automation framework it is not about the operator taking over tasks from the automation, but adapting towards a more suitable collaboration style for the situation. The conclusion based on the user evaluation is that DP-operators with extensive experience with the traditional way of working including working with traditional conning displays, is that the joint human-automation framework is rated as positive, that is of added value and definitely has potential.

The fact that we were able to conduct a user evaluation underlines the importance of applying use cases for knowledge development in early research projects. Despite the fact that the focus of the project was not on solving use case specific issues, the application of a use case was instrumental in demonstrating the added value of a joint human-automation framework. In order to be able to simulate realistic, i.e. for DP-operators recognizable, DP-scenarios, it was necessary for the researchers to understand the DP-practice on a detailed level. Generally speaking, a human centered system development approach does not only require human factors knowledge but also extensive domain knowledge, especially knowledge about the operational variability and challenges to which operators are exposed.
In addition to the fact that we were able to demonstrate the added value of technology we were also able to demonstrate that a joint human-machine framework enables another way of working.

Another aspect of use case demonstration is that it is necessary to demonstrate the applied nature of cutting edge research and, hence, demonstrate the added value and market potential. The difficulty about this is, that at the start of the early research program the use case approach was also translated, by the management, into valorization targets to be met during the project. The conclusion is that pursuing both targets--develop cutting edge knowledge and techniques and to sell them as well—at the same time proved not to be possible. To bring cutting edge knowledge and techniques to the market involves a secondary market development effort of finding market partners, setting up pilot studies and implementation using first adopters.

The core of a joint human-automation framework is the ability to adapt towards a more suitable collaboration style for the situation. This research shows that the Intelligent Operator Support System provides the support needed for adapting towards suitable collaboration styles for different situations. The concentration (vigilance) that monitoring of highly automated autonomous systems requires, is, in general, difficult for people to uphold. This could lead to performance problems that operator situation awareness (SA) is not sufficient to be able to intervene effectively. The vigilance problem was the reason to change the collaboration style. The IOSS takes over the monitoring task in situations that are stable internally (the DP system) and externally (environmental conditions), and provides SA recovery support by bringing the operator back into the loop. This approach shows that by changing several fundamental aspects of the system that determine how the human and automation will interact, how roles and responsibilities will be allocated between them, and how often these allocations will change, has a significant effect on the complexity of the system and the level of engagement and workload of the operator, all of which significantly influence system oversight and intervention.

This kind of adaptability towards more suitable collaboration styles for different situations cannot be achieved with more traditional automation and interfaces, but, instead, requires a different, intelligent form of automation and advanced ways of interaction. The IOSS is a so-called intelligent agent since its functionality is dedicated to maintain operator SA and to support SA recovery when needed. Central to the notion of agency is the *human aware* ability of technical systems. It is the ability of technical systems to create a model of the operator state and to adjust operator support according to the operator state. What operator state involves, depends on the domain, task and situation. In cars information about whether a driver is getting sleepy or does not pay attention to the road, could be the trigger to take action to restore driver involvement or to give the advice to take a coffee break. In the DP use case the human aware capability was confined to detecting whether or not the operator was roaming on not.

However, two experiments conducted within the Adaptive Maritime Automation project give rise to the conclusion that more elaborate applications of human aware computing AI are possible. One experiment involved the modeling of the operator cognitive state during several training sessions in a DP-simulator. The way the trainees handle different failure modes was captured and compared with the expert solution of the DP-trainer. The outcome was that most errors were not the result of failures going undetected, but the result of the fact that operators were unable to come up with the correct courses of action to mitigate the situations. These findings illustrate that it is potentially beneficial to extend the functionality of the IOSS with decision making support. Hence, to provide support in situations in which operators are being unable to come up with the correct courses of action to mitigate the situation.

A second experiment did confirm the detrimental effects of ill-designed support on performance: The static and the critical event support conditions, both causing an abundance of alarms, most of them false, decreased overall system performance. So, the conclusion could be that hybrid adaptive support prevented the operator to become overloaded with alarms, thereby allowing him or her to maintain stable system performance on the relatively short duration of the task. This could be interpreted as support for the concept of smart notifications.

The smart notification concept entails information push (instead of information pull as is the current practice) as well as a set of design patterns for establishing interaction between humans and highly autonomous systems.
In the future, maritime operations will become more complex and internet of things approaches are expected to produce massive amounts of data. Also, cooperation between people and autonomous systems will increasingly take place over a distance via cyber networks in so-called cyber-physical systems. The question then is whether the remote human operator, isolated from the reality of the autonomous system (ship, drone etc.) and its actual environment, can technically be supported in such a way that this person is sufficiently able to secure the autonomous system and to achieve operational objectives. In light of these developments, the conclusion is justified that intelligent operator support will become even more important in the future and that more human factors research and development are justified.

7 References

Arciszewski, H. F., De Greef, T. E., and Van Delft, J. H. (2009). Adaptive automation in a naval combat management system. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 39(6), 1188-1199.

Bradshaw, J. M., Hoffman, R. R., Woods, D. D., & Johnson, M. (2013). The Seven Deadly Myths of" Autonomous Systems". IEEE Intelligent Systems, (3), 54-61.

Bradshaw, J. M., Feltovich, P., Johnson, M., Breedy, M., Bunch, L., Eskridge, T., Jung, H., Lott, J., Uszok, A., van Diggelen, J. (2009, July). From tools to teammates: Joint activity in human-agent-robot teams. In International Conference on Human Centered Design (pp. 935-944). Springer Berlin Heidelberg.

Broek, J. van den (2017). Techniek is belangrijk, maar het zijn mensen die het verschil maken: De relevantie van human factors in maritieme automatisering. [openbare les]. Hogeschool Rotterdam Uitgeverij. www.hr.nl/onderzoek/publicaties.

Broek, J. van den, Schraagen, J.M.C, Brake, G.M. te & Diggelen, J. van (2017). Approaching full autonomy in the maritime domain: paradigm choices and Human Factors challenges. Proceedings of MTEC2017, 26-28 april 2017, pp. 375-389. Singapore.

Bunch, L., Breedy, M., Bradshaw, J. M., Carvalho, M., Danks, D., Suri, N. (2005, March). Flexible automated monitoring and notification for complex processes using KARMEN. In Proceedings. 2005 IEEE Networking, Sensing and Control, 2005. (pp. 443-448). IEEE.

Endsley, M. R. (2016). From Here to Autonomy: Lessons Learned From Human–Automation Research. Human Factors, 0018720816681350.

Endsley, M. R., Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 42, 462–492.

Endsley, M. R., & Kiris, E. O. (1995). The outof-the-loop performance problem and level of control in automation. Human Factors, 37(2), 381-394.

Forehand, M. (2010). Bloom's taxonomy. Emerging perspectives on learning, teaching, and technology 41 (2010), 47.

Gunning, D. (2017) Explainable artificial intelligence (XAI). Defense Advanced Research Projects Agency (DARPA) (2017).

Hockey, G.R.J. (1997). Compensatory control in the regulation of human performance under stress and high workload: a cognitive energetical framework. Biological Psychology, 45, 73–93.

Heinen, G.A. (2016). Enhancing safety and reliability in Dynamic Positioning: The role of Situational Awareness for Dynamic Positioning Operators performance. (Thesis). University of applied sciences Rotterdam.

Houtkoop, K. (2017). Effectively monitoring problem solving in Dynamic Positioning. (Thesis). University of applied sciences Rotterdam.

Jasper van der Waa, van Diggelen Jurriaan, Mark Neerincx, and Stephan Raaijmakers (2018) ICM: An intuitive, model independent and accuracy certainty measure for machine learning. In 10th Int. Conf. On Agents and AI.

Kaber, D. B. (2013). Adaptive automation. In J.D. Lee & A. Kirlik (Eds.), The Oxford Handbook of Cognitive Engineering (pp. 594-609). New York: Oxford University Press.

Kaber, D. B., & Endsley, M. R. (1997). Out-ofthe-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. Process Safety Progress, 16(3), 126-131.

Kingma, D., Ba, J. (2015). Adam: A method for stochastic optimization. 3rd Int. Conf. for Learning Representations.

Lee, H. G. (2013). A study on predictive analytics application to ship machinery maintenance (Doctoral dissertation, Monterey California. Naval Postgraduate School).

Matthews, G., Davies, D.R., Westerman, S.J. & Stammers, R.B. (2000). Human Performance: Cognition, Stress and Individual Differences. London: Psychology Press.

Meyer-Bisch, C. (2005). Hypoacousie due au bruit: la réglementation évolue. Médecine et Sciences, 21 (12), 1089–1095.

Miller, T. (2017). Explanation in artificial intelligence: Insights from the social sciences. arXiv preprint arXiv:1706.07269 (2017).

Nair, V., Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proc. Of the 27th Int. Conf. on Machine Learning (ICML-10), pages 807–814.

Neerincx, M. A., and Lindenberg, J. (2008). Situated cognitive engineering for complex task environments. In J.M. Schraagen, L.G. Militello, T. Ormerod, and R. Lipshitz (Eds.), Naturalistic Decision Making and Macrocognition (pp. 373-389). Aldershot: Ashgate Publishing Limited.

O’Donnell, C. R., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of perception and human performance: Vol. II. Cognitive processes and performance (pp. 42.1-42.29). New York: Wiley.

Parasuraman, R., Cosenzo, K. A., de Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. Military Psychology, 21(2), 270.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Poelman, M. (2017). Cognitive state of Dynamic Positioning Operators. (Thesis). University of Apllied Sciences, Rotterdam.

Sarter, N.B., Woods, D.D. & Billings, C.E. (1994). Automation surprises. In: G. Salvendy, Handbook of human factors and ergonomics (pp. 1926–1943). John Wiley and Sons.

Sauer, J. Chung-Shan, K., Wastell, D., Nickel, P. (2011). Explicit control of adaptive automation under different levels of environmental stress, Ergonomics, 54:8, 755-766, DOI: 10.1080/00140139.2011.592606.

Sauer, J., Nickel, P., Wastell, D. (2013). Designing automation for complex work environments under different levels of stress. Applied ergonomics, 44(1), 119-127.

Sheridan, T. B., Verplank, W. L. (1978). Human and computer control of undersea teleoperators. MASSACHUSETTS INST OF TECH CAMBRIDGE MAN-MACHINE SYSTEMS LAB.

Singh, H., Bhatia, J. S., Kaur, J. (2011). Eye tracking based driver fatigue monitoring and warning system. In Power Electronics (IICPE), 2010 India International Conference on (pp. 1-6). IEEE.

Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircraft systems design. In Situational awareness in aerospace operations (Rep. No. AGARDCP-478) (pp. 341–346). Neuilly Sur Seine, France: NATO-AGARD.

TNO. (2016). Human Enhancement by Maritime Adaptive Automation. Video, TNO publication (https://youtu.be/MH0Vj-rChrM), TNO Soesterberg.

TNO. (2017). Adaptive Maritime Automation. Chapter in: *Human Enhancement*. TNO Brochure, 2017.

Van der Kleij, R., Hueting, T., & Schraagen, J.M. (2018). Change detection support for supervisory controllers of highly automated systems: Effects on performance, mental workload, and recovery of situation awareness following interruptions. International Journal of Industrial Ergonomics, 66, 75-84. doi.org/10.1016/j.ergon.2018.02.010.

Van Lent, M., Fisher, W., Mancuso, M. (2004). An explainable artificial intelligence system for small-unit tactical behavior. In Proceedings of the National Conference on Artificial Intelligence (pp. 900-907). Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press.

Veltman, J. A., Gaillard, A. W. K. (1993). Measurement of pilot workload with subjective and physiological techniques. Paper presented at the Royal Aeronautical Society on the assessment of workload and aviation safety.

Zijlstra, F. R. H. (1993). Efficiency in work behaviour: A design approach for modern tools. Doctoral dissertation, Delft University of Technology, Delft, The Netherlands.

A Raw interview data

A.1 Scenario 1

A.1.1 Notification 1

Vraag	2 Begrijpt u notificatie?
PP1	
PP2	I would like a fixed scale for the sensor-value graphs
PP3	duidelijk
PP4	It is clear, but I would just like one checkbox to notify me that everything is okay [instead of having to monitor all values]
PP5	Unclear what is meant with ‘changes’, is that peak values? Gradual increase? Superfluous, I can already see that in the DP system [at this point operator was not roaming]

Vraag	3 opvolgend?
PP1	ja
PP2	
PP3	Yes, especially because i can see relevant information, in the beginning you will check this often but will rely on it after more experience and trust
PP4	Not until i have trust in the system
PP5	No, certainty (85%) is too low

Vraag	4 nut zekerheid?
PP1	Hangt af van of er op tijd gewaarschuwd wordt, zekerheid moet wel boven 75% zijn. Advies moet **niet** gegeven worden als zekerheid laag is
PP2	
PP3	Goed to know
PP4	I don’t know what ‘85%’ means [what is it based on? (see ICM)]
PP5	yes

Vraag	5 meer informatie?
PP1	The weather forecast
PP2	The variance of the sensor values
PP3	I would like ONE button to show me ALL information, [instead of having to open one value at a time] Graphs could be made easier to interpret
PP4	Some additional sensors
PP5	I want to know where the insecurity comes from, and i want the system to summarise whether everything is under control or not.

Vraag	6 bent u nodig?
PP1	No, i’ll hear it when I am
PP2	
PP3	
PP4	
PP5	

A.1.2 Notification 2

Vraag	2 begrijpt u?
PP1	I don’t expect an alarm until the DP system can’t hold position anymore
PP2	yes
PP3	yes
PP4	yes
PP5	-

Vraag	3 opvolgen?
PP1	yes
PP2	Yes
PP3	yes
PP4	yes
PP5	no

Vraag	4 nut zekerheid?
PP1	Het is onduidelijk, liever gekwantificeerd in %
PP2	No, advice to return did it, regardless of 40/80% certainty
PP3	No, but influences feeling of urgency
PP4	No, i would go and have a look anyway, if it’s something serious you have to be there early
PP5	No, you want to be safe in any case

Vraag	5 reden voor terugkeer duidelijk?
PP1	ja, toen ik op de popup klikte zag ik wind toename
PP2	yes
PP3	Yes, but i would like to see what it is that is expected
PP4	yes
PP5	yes

Vraag	6 meer informatie?
PP1	Waar is advies op gebaseerd? Ik wil advies telkens opnieuw krijgen
PP2	All buttons give roughly the same information, i would rather have one single ‘more info’ button Also variance of reference systems
PP3	-
PP4	What ARE the changes?
PP5	‘ik wil feedback dat mens weer terug’

A.1.3 Notification 3

Vraag	1 begrijpelijk?
PP1	Waarom vind je dat ik moet blijven zitten, wat weet jij wat ik niet weet?
PP2	-
PP3	clear
PP4	No, i thought the message meant everything was okay but it was not
PP5	No, if it is said that the situation will remain the same, it implies that you can leave

Vraag	2 opvolgen?
PP1	Als er kans is dat het misgaat blijf ik zitten
PP2	-
PP3	I would stay in the area, but maybe not behind the desk, I assume IOSS would warn me in time.
PP4	yes
PP5	yes

Vraag	3 nut zekerheid?
PP1	Nee, ik wil de limieten weten
PP2	-
PP3	No, i will stay with the DP desk regardless
PP4	With low certainty i don't trust the calculations display certainty like a 'cell phone reception graph'
PP5	I don't want a value, but a reason [for less certainty]

Vraag	4 meer informatie?
PP1	Vermogen van de individuele schroeven
PP2	-
PP3	No, if IOSS says it will remain the same, there is no need for me to look
PP4	I would like to know the predections
PP5	What the prediction is based on

A.1.4 New working agreements

Vraag	1 make new working agreements?
PP1	-
PP2	
PP3	
PP4	Wil kunnen opvragen een per dag aanpassen
PP5	Make working agreements about Intuitive Certainty Measure, instead of having ICM in the notification

A.2 Scenario 2

A.2.1 Notification 1

Vraag	3 begrijpelijkheid
PP1	-
PP2	-
PP3	Very clear
PP4	yes
PP5	Again, a % i don't want to see

Vraag	4 opvolgen
PP1	Ja, geen verdere informatie [nodig?]
PP2	yes
PP3	Yes, I think so
PP4	Yes, provided there is trust in the system
PP5	Not at this stage, maybe after i've had some more experience with IOSS

Vraag	5 nut zekerheid
PP1	ja
PP2	A percentage is enough, elaboration on how it got to percentage is superfluous
PP3	Yes, it’s definitely important
PP4	Yes, very good, i don’t know about %, rather have something like 4/5 ‘green bars’ = good
PP5	yes

Vraag	6 meer informatie nodig
PP1	Nee, windsnelheid is voldoende
PP2	no
PP3	Not in this situation
PP4	No, i already know what the situation is (operator is not roaming)
PP5	No, i already know what the situation is (operator is not roaming)

Vraag	7 bent u nodig?
PP1	Nee, hij heeft me **verzekerd** dat ik weg kan, hij zei niet dat hij me waarschuwt maar daar ga ik vanuit
PP2	-
PP3	-
PP4	-
PP5	-

A.2.2 Notification 2

Vraag	1 begrijpelijk?
PP1	Er lijkt niks aan de hand, maar waarom alarm (niks aan de hand = schip ligt nog in positie)
PP2	[gaat gelijk terug, pakt daarna IOSS erbij]
PP3	ja
PP4	ja
PP5	-

Vraag	2 opvolgen?
PP1	Ja, zeker
PP2	ja
PP3	ja
PP4	Ja, ga meteen terug
PP5	Ja, sowieso terug

Vraag	3 reden terugkeer duidelijk?
PP1	Ja
PP2	Ja, gelijk terug
PP3	ja
PP4	ja
PP5	Ja, maar niet precies genoeg, [er staat wat het nu is, maar niet wat het zou moeten zijn]

Vraag	4 meer informatie nodig
PP1	Power & richting schroef, (+positie & vermogen andere thrusters (zit in widget))
PP2	Ik zou alle thrusters afzonderlijk willen zien ik zou ook trend data willen kunnen zien
PP3	Maakt niet uit [ik ga sowieso terug], kijk al lopend IOSS alarm moet in dezelfde taal als DP systeem
PP4	Ik wil liever een stoplicht [oranje/rood = terug]
PP5	Ik wil afwijking t.o.v. normaal zien ik wil ook het directe gevolg voor position-keeping [doel] zien.

A.2.3 Notification 3

Vraag	1 begrijpelijk
PP1	-
PP2	-
PP3	Misschien overbodig
PP4	Overbodig
PP5	Overbodig, maar kan geen kwaad

Vraag	2 opvolgen
PP1	Ja, Ik blijf zitten, is gevoelsmatig wat ik zelf ook wilde
PP2	Afhankelijk van operatie, in dit geval overbodig [om te vermelden]
PP3	Ja, zeker
PP4	ja
PP5	ja

Vraag	3 meer informatie?
PP1	Wil de thruster waarden weten, maar ik zit nu al achter de DP console Had de melding niet hoeven krijgen, maar fijn dat IOSS bevestigt
PP2	-
PP3	Nee, je weet het al [operator zit achter DP desk]
PP4	Informatie is summier, er staat bijv. niet welke thruster
PP5	Nee, duidelijk genoeg

A.2.4 New working agreements

Vraag	1 make new working agreements?
PP1	Password on working agreements to prevent accidental changes work with operator profiles [i.e. personalised settings]
PP2	-
PP3	No unless you plan to continue with one failed thruster, then you need new (lower) limits
PP4	Settings for feedback error [i.e. difference between desired & actual]
PP5	Make working agreements about Intuitive Certainty Measure, instead of having ICM in the notification

A.3 Usefulness functions

functie	Pp1	Pp2	Pp3	Pp4	Pp5
DP status scherm.	V	X	V	V	V
De avatar die de algemene stress van IOSS en het DP systeem vertegenwoordigd. [zat er nu niet in]	X, word je zenuwachtig van	X	X	V	V
De mogelijkheid om meer informatie op te vragen via de notificaties.	V	X	X	V	X
Notificaties van de DP alarmen.	V	V	V	V	V
Notificaties met het advies dat u uw werkplek kunt verlaten of juist terug moet keren.	V	X	V	V	V
De grafieken in de smart notificaties.	V	V (wil historie van paar minuten kunnen zien)	V (mits constante schaal)	V	X
De notificaties over DP alarmen.	V	V	V	X	V
De zekerheid of het advies dat IOSS geeft inderdaad klopt.	V (als je weet wat erachter zit hoef ik % niet te zien)	V	V	V	V (maar simpeler weergegeven)
Uitleg hoe IOSS op het advies om uw werkplek te verlaten of juist terug te komen is gekomen.	V	V	V	X	V
De mogelijkheid om zelf afspraken te maken met IOSS over wanneer welke notificaties gegeven worden.	X	X	X	X	X
De checklists die IOSS kan afnemen die zich aanpassen aan welke informatie u wel of niet heeft gezien of eventueel zijn veranderd.	X		X	X	X

functie	Pp1	Pp2	Pp3	Pp4	Pp5
De mogelijkheid van IOSS om automatisch te detecteren of u wel of niet bij uw werkplek in de buurt bent.	X	X	V (belangrijk)		X

A.4 Trust

Vraag	1 Would wrong advice influence your trust, what if IOSS already warned about low certainty?
PP1	nee
PP2	Yes, unless low certainty was already given
PP3	yes
PP4	yes
PP5	Certainty measure is useful for making a well-thought out decision

Vraag	2 Are you afraid you will miss information?
PP1	Nee, goed instellen = goede output
PP2	I need to be able to see relevant information
PP3	no
PP4	yes
PP5	no

Vraag	3 if IOSS is uncertain, do you still want advice?
PP1	'15 min advies blijven zitten' [bij notificatie 1?]
PP2	Niet zinvol, dan blijf ik wel zitten
PP3	Warnings, and 'return to desk'-notification yes, but not advice to leave desk (not until +/-80% certain)
PP4	Useless, except for calling back to DP desk
PP5	Yes, but with the warning that IOSS is not sure.

A.5 Usefulness IOSS

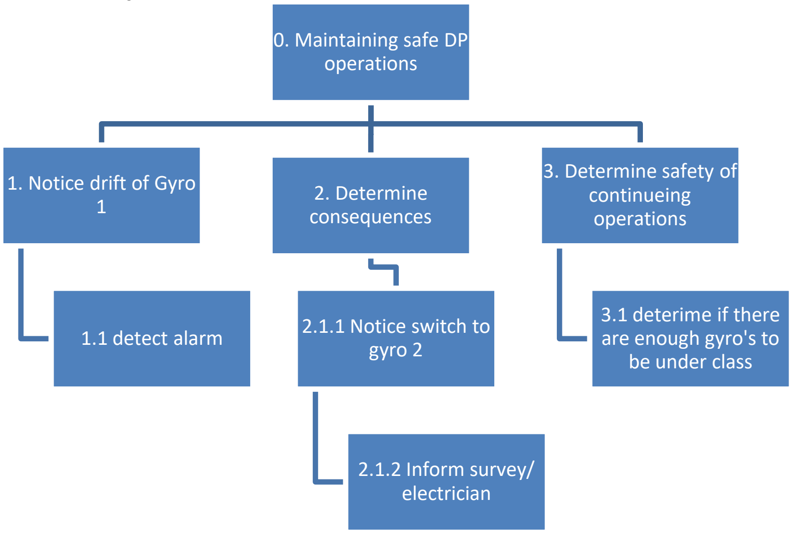
Vraag	1
PP1	Ja, 8, maar er zullen nog een hoop heilige huisjes om moeten
PP2	Ja, mits operatie het toelaat 6-7 rock dumping, 9 stable operations
PP3	Yes, 8, depending on kind of operation
PP4	Yes, so that captain can also monitor remotely, for operator assistance, and for centralising information
PP5	

B Optimal solution path for each failure mode

Descriptions in the appendix by courtesy of Houtkoop, K.C.F. (2017).

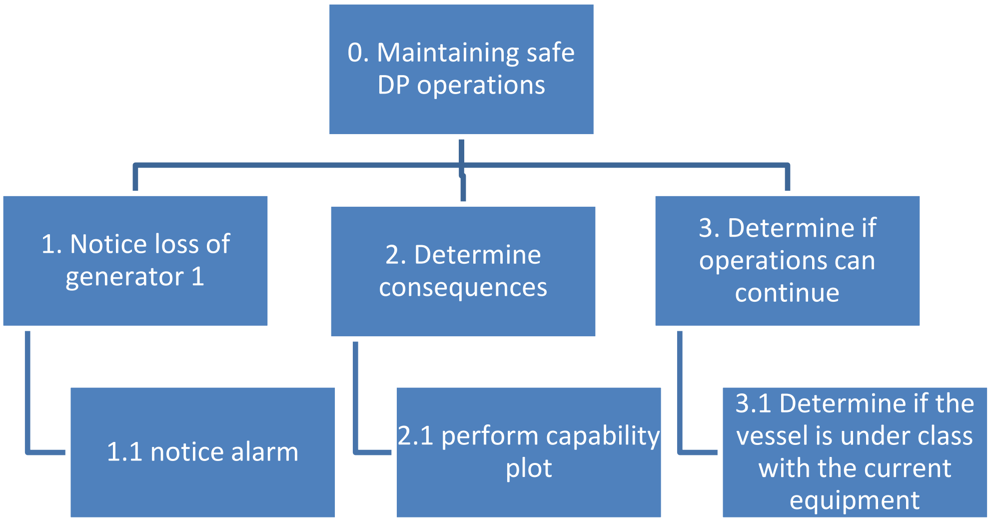
B.1 Day 1

B.1.1 Failure mode 1 – Gyro drift



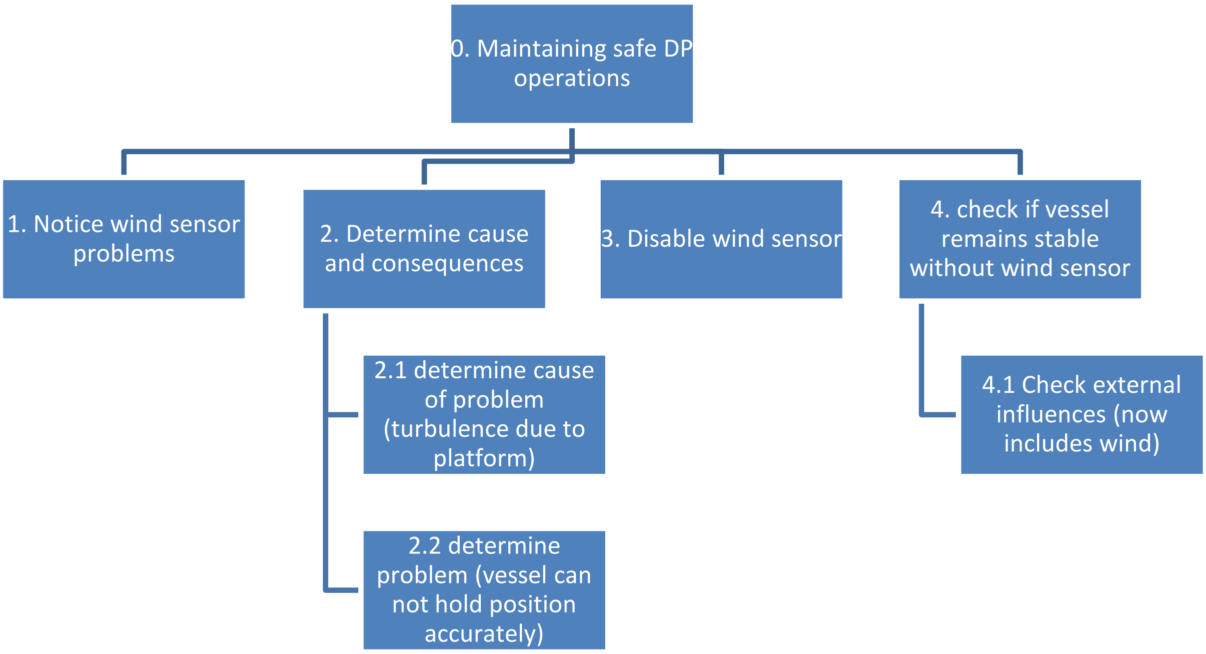
Place in HTA	Information/action required	Observable action
1>1.1	Notice drift of gyro	•looking at alarm
2>2.1	Observe status of sensor equipment	•Look at sensor page
2.1>2.1.1	Notice the system automatically switched to Gyro 2	•Notice switch on sensor page
2.1>2.1.2	Inform survey/electrician	•Call survey/electrician by phone
3>3.1	Determine if operations can safely continue. (3 gyro's are required to be under class)	•Continuing or stopping of operations shows decision

B.1.2 Failure mode 2 – Generator failure



Place in HTA	Information/action required	Observable action
1>1.1	Notice loss of generator 1	•see alarm • notice generator failure on power screen
2>2.1	Determine consequences for vessel capabilities	•Performing capability plot •Check available power
3>3.1	Determine if vessel can continue operations under class	•Discuss if vessel is under class •Continue operations •Notice there are **no** consequence alarms

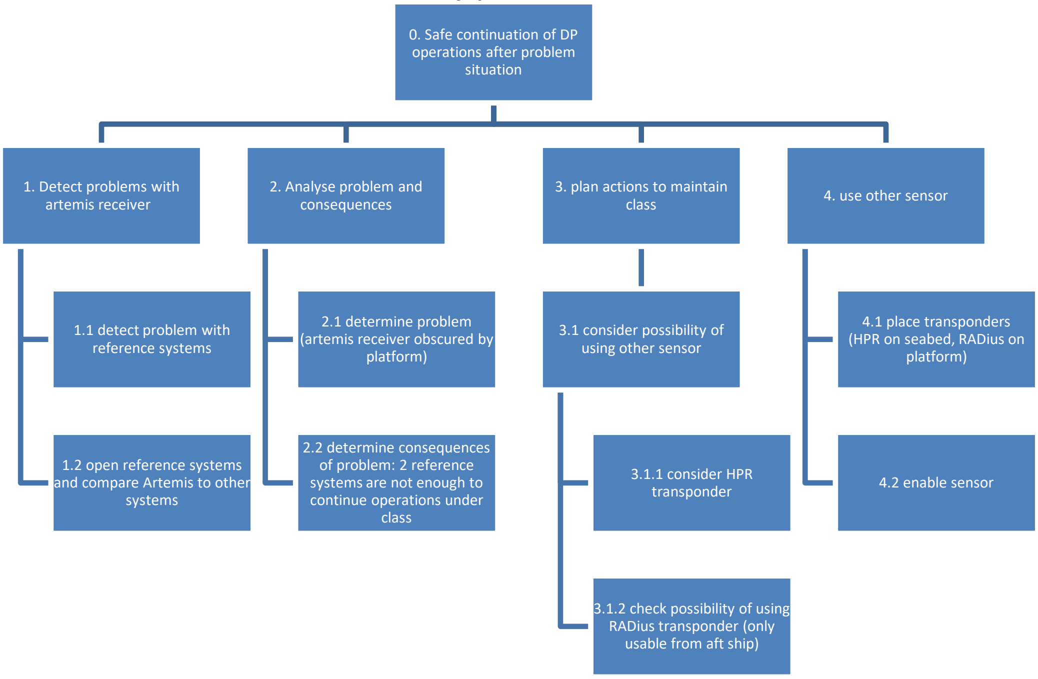
B.1.3 Failure mode 3 – unstable wind sensor



Place in HTA	Information/action required	Observable action
1	Notice wind sensor problems	•Looking at wind direction (main overview)
2>2.1	Determine cause of problem (turbulence)	•Wind speed and direction plot •**<u>Wrongly</u>** determined if they send electrician to sensor
2>2.2	Vessels position accuracy is lowered	•Notice by looking outside (ship moving relative to platform) •look at drift on main overview
3	Disable wind sensor	•Disabling the wind sensor at the sensor page
4>4.1	Vessel is stable and wind influence is now compensated as current	•Wind arrow no longer available on main overview •Current direction is stronger/weaker dependent on wind.

B.2 Day 2

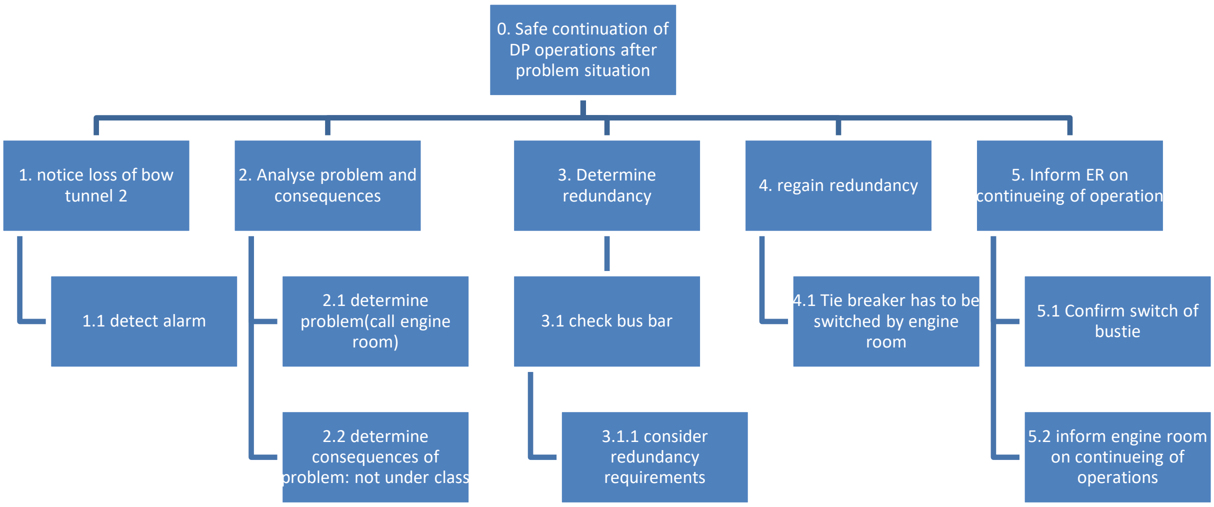
B.2.1 Failure mode 1 – Artemis obscured by platform



Place in HTA	Information/action required	Observable action
1>1.1	Detect problems with reference systems	•Look at available systems on reference systems page
1>1.2	Detect that the artemis is causing problems	•Notice that the artemis accuracy is going down •Notice that the artemis is cutting out
2>2.1	Determine problem (artemis is obscured)	•Can be planned beforehand, actions will be taken before it happens (no obstruction occurs but other steps will be done). •Look at chart of operation and notice obstruction •Looking outside and noticing obstruction
2>2.2	Determine consequence (2 reference systems are not enough to be under class)	1. start using other sensor (see next step) 2. informing survey about use of other sensor 3. pause of approach

Place in HTA	Information/action required	Observable action
3>3.1	Two sensors are not enough to maintain class and continue. Another one has to be used.	•Looking at sensors available on sensor screen •Call survey about other sensors •Pause approach
3.1>3.1.1	Consider HPR transponder	•Using the HPR transponder by calling survey and deploying
3.1>3.1.2	Consider possibility of RADius transponder	•Using Radius transponder by calling platform and deploying (only possible when entering stern first)

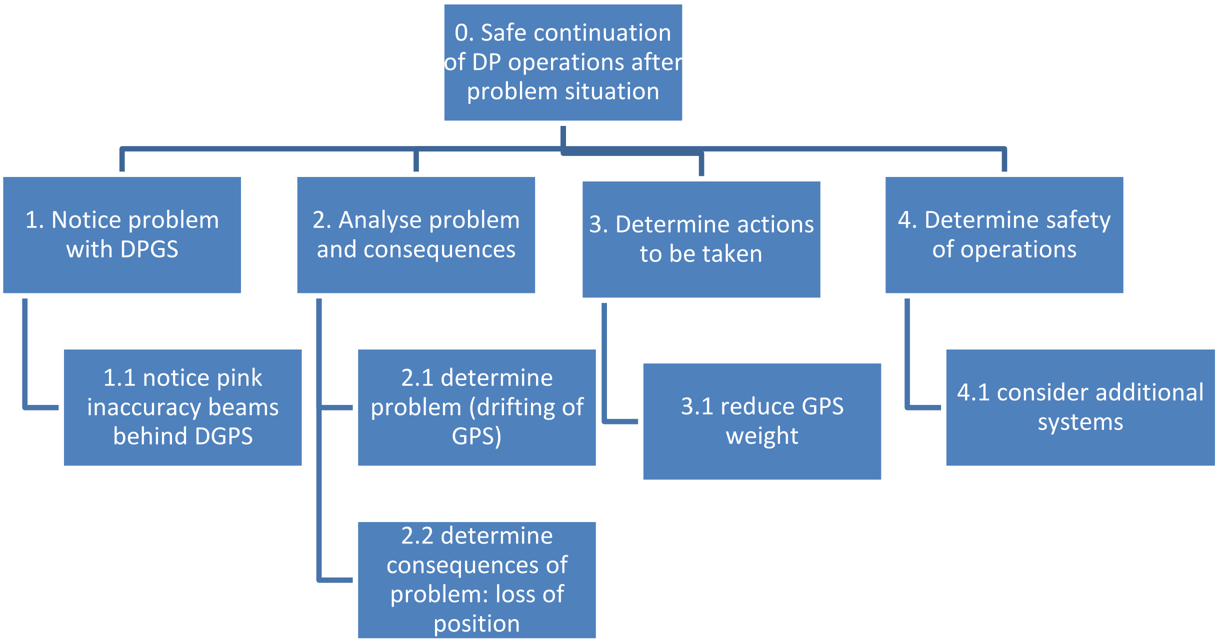
B.2.2 Failure mode 2 – Loss of bow-tunnel thruster



Place in HTA	Information/action required	Visible action
1>1.1	Notice loss of bow tunnel 2	•Look at alarm list •Look at thruster overview page
2>2.1	Determine problem	•Same as previous •calling of engine room
2>2.2	Determine consequences for operations	• Making of capability plot • Notice consequence alarm
3>3.1	Check bus bar for redundancy	•Look at busbar on power overview screen
3>3.1.1	Notice that bow thrusters are not on separate busbars	•Look at tie breaker

Place in HTA	Information/action required	Visible action
4>4.1	To regain redundancy the engine room has to switch the tie breaker	•Engine room called to switch •Tie breaker switches due to engine room
5>5.1	Confirm switch of bustie to redundant system	•Look at power overview bustie breaker (breaker switched to other busbar)
5>5.2	Call engine room confirm continuing of operations	•Use of phone

B.2.3 Failure mode 3 – DGPS drift



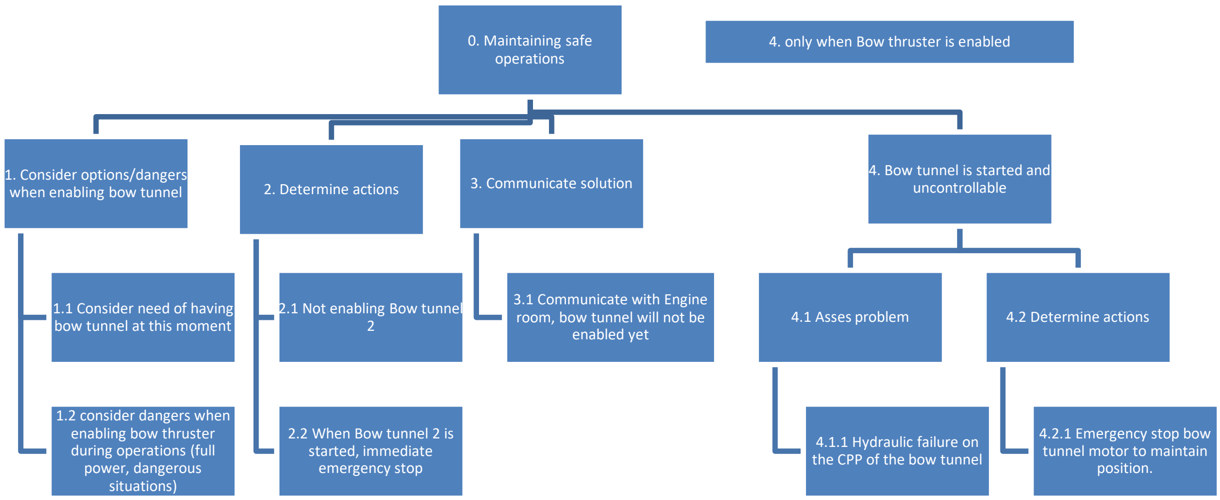
Place in HTA	Information/action required	Observable action
1>1.1	Notice GPS problem	•Notice pink inaccuracy bars at the reference systems (top left)
2>2.1	Determine problem (DGPS drift)	•Notice both DGPS differ from 3rd system (reference systems page) •**<u>Wrong</u>** if the 3rd system is considered wrong
2>2.2	Determine consequences for operation	•Notice that there is only 1 system available (sensors page)
3>3.1	Determine actions to be taken (reducing of GPS weight)	•Reduce the weight of the GPS at the sensor screen.

Place in HTA	Information/action required	Observable action
4>4.1	Determine if the operation can continue	•Not necessary but additional systems can be used (enable at sensors) •Survey can be informed to place additional systems

Comment from DP instructor:

"Reduced weight should have been enabled from the start, if this problem occurs then there was a problem at the setup."

B.2.4 Failure mode 4 – Re-engaging bow-tunnel thruster after faulty repair



Place in HTA	Information/action required	Observable action
1>1.1	Consider need of having the bow tunnel at this moment	•Capability plot •Looking at available power
1>1.2	Consider dangers when enabling tunnel	Not visible, visible due to following actions taken by the operator.
2>2.1	Not enabling bow tunnel	•Telling the engine room not to enable the bow tunnel
3>3.1	Communicate with engine room	•Communicate with engine to not start the bow tunnel

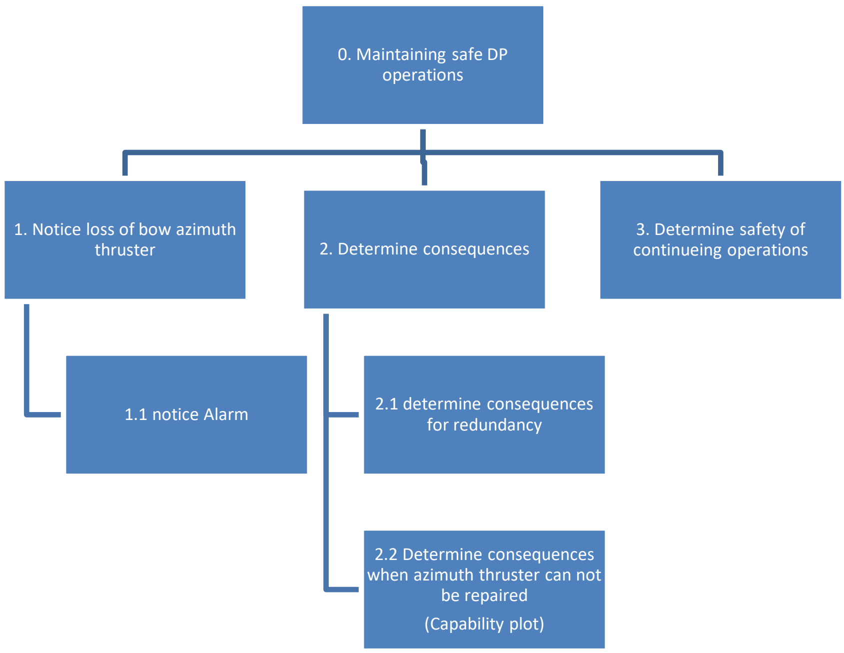
Place in HTA	Information/action required	Observable action
4 >4.1	Asses problem why bow tunnel is uncontrollable	•Notice movement of ship
4.1>4.1.1	Bow tunnel gives no pitch on screen but is moving the ship. This indicates a problem in the CPP	•on propulsion screen notice that CPP gives no pitch but vessel is moving
4>4.2	Determine actions to stop the uncontrollable CPP	Not visible, see next step
4.2>4.2.1	Stopping the CPP motor to stop the influence on the ship	•Stop the bow thruster using emergency stop •**Wrong** if only deselected at the DP console (operator does not recognize problem)

Comment from DP instructor:

"This is a situation where the mistake is to enable or test equipment during an operation, especially on critical stages of the operations like in this instance: the divers. It would be difficult for a system like the IOSS to assist here, this is something that should be impossible to do."

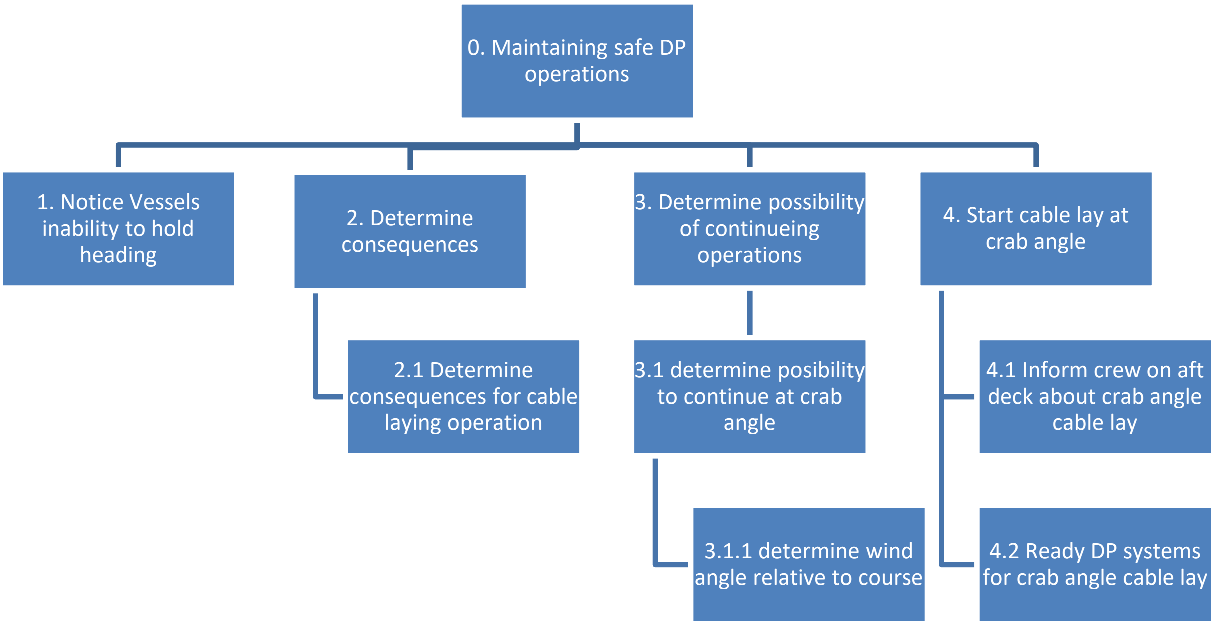
B.3 Day 3

B.3.1 Failure mode 1 – bow-azimuth thruster failure



Place in HTA	Information/action required	Observable action
1>1.1	Notice loss of bow azimuth thruster	•Notice alarm •Notice thruster not working on thruster screen
2>2.1	Determine consequences for redundancy	•Look at busbar (tie breaker position) •Look at thruster screen
2>2.2	Determine consequences when azimuth thruster cannot be repaired	•Making of capability plot
3	Determine if operations continue	•Either continuing or stopping operation

B.3.2 Failure mode 2 – wind exceeding limits



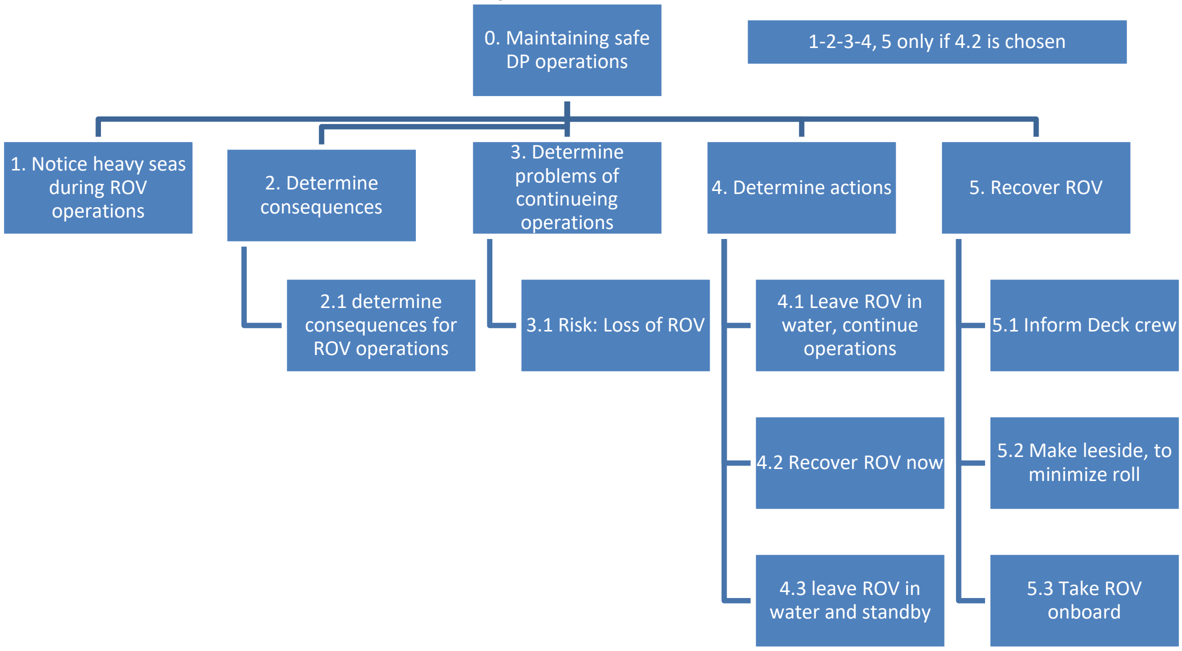
Place in HTA	Information/action required	Observable action
1	Notice vessel cannot hold heading	•Notice vessel cannot hold heading •Notice heading out of limits alarm
2>2.1	Determine consequences for cable lay operation (30 degrees max)	•Continuing if heading stays within 30 degrees of cable-lay •Look at power overview screen (power usage is increasing)
3>3.1	Determine if the vessel has to continue at crab angle	•Compare heading and wind direction (look at overview screen) •Look at power overview (power increasing)

Place in HTA	Information/action required	Observable action
3.1>3.1.1	Determine crab angle (max 30)	•Check wind direction (for a more head on wind)
4>4.1	Inform crew on aft deck about cable lay at crab angle	•Call crew on aft deck
4>4.2	Ready system for cable lay	•Change heading angle to 30* from course manually (user control) •Change heading using crab angle function •Change heading in track list

Comment from DP instructor:

> *"Situations like these require early action, and preferably more information. The weather is difficult to predict and this should be done with the means available. The choice to continue the operation can be a choice of safety, however it can also in some cases be a financial chose: where the gain outweighs the risk of damage (within reason)."*

B.3.3 Failure mode 3 – wind exceeding ROV limits



Place in HTA	Information/action required	Observable action
1	Notice heavy seas	•Look at weather trend screen •Look outside •Notice vessel rolling
2>2.1	Determine consequences for ROV operations	•Notice that rolling is increasing in weather trend. Following the next steps means that the operator realizes there are consequences
3>3.1	Risk of loss ROV	•When ROV is recovered in the next steps this is mitigated. •When left in the water this risk is accepted.
4>4.1	Leave ROV in water	•Continuing operations
4>4.2	Recover ROV now	•Informing ROV crew and deck crew that ROV will be recovered
4>4.3	Leave ROV in water and standby	•Informing ROV crew that ROV will be on standby until better weather
5>5.1	Inform deck crew about recovery	•Calling deck crew and informing them about recovery
5>5.2	Make leeside to minimize roll	•Change of heading to make leeside
5>5.3	Take ROV onboard	•Deck crew will inform when ROV is onboard

Comment from DP instructor:

“When the weather is becoming worse the choice to continue or stop the ROV operation has to be made: this is done according to information known beforehand like a maximum wave height or wind speed. This is a choice to be made together with the captain and the company supervisor most of the time.”