



Försvarshögskolan

Självständigt arbete (30 hp)

Författare		Program/Kurs
Marcus Johansson		HOP SA 2024
Antal ord: 9994		
(varav 163 i bilder)		
Handledare		Kurskod
Maria Eriksson Baaz		2UK023
<p>Leva eller dö - ett beslut av människa eller maskin?</p> <p>En deskriptiv studie om artificiell intelligens i militära operationer.</p>		
<p>ABSTRACT:</p> <p>Studies on using artificial intelligence in a military context often focus on either technical aspects or legal and ethical perspectives. There is, however, a gap in the current academic AI debate regarding one of the key components in all military operations: The officers in charge.</p> <p>This study aims to analyze how Swedish officers think and perceive the introduction of AI in decision-making in a military context. How do they reason, in favor or against the use of AI on the battlefield?</p> <p>A novel method within war science, the vignette survey, will be used to analyze how human-AI relationships are affected by an increased level of risk of loss of human life in three different vignettes. The findings in this study describe that officers trust AI to be used in military operations, regardless of the complexity or risk of the operation. As long as a human is in command.</p>		
<p>Nyckelord:</p> <p>Artificiell intelligens, Beslutsfattning, Tillit, Militära operationer.</p>		

Innehåll

1. INLEDNING	3
1.1. BAKGRUND:	3
1.2. SYFTE.....	4
1.3. DISPOSITION.....	4
1.4. PROBLEMSTÄLLNING OCH FORSKNINGSFRÅGOR.....	5
1.5. AVGRÄNSNINGAR:	5
1.6. EXTERNA SAMARBETEN:	6
2. FORSKNINGSÖVERSIKT	7
3. TEORI.....	11
3.1. CENTRALA BEGREPP.....	11
3.2. FÖRVÄNTNINGAR PÅ RESULTATET UTIFRÅN TIDIGARE FORSKNING:	12
4. METOD	16
4.1. URVAL	16
4.2. FORSKNINGSETIK OCH RESPONDENTER	16
4.3. DESIGN.....	18
4.4. OPERATIONALISERING	20
5. RESULTAT	23
5.1. VINJETT 1	23
5.2. VINJETT 2.	25
5.3. VINJETT 3	27
6. ANALYS OCH DISKUSSION:.....	28
6.1. TILLIT TILL AI I MILITÄRA OPERATIONER.....	28
6.2. BIDRAG TILL PROFESSIONEN	31
6.3. METODOLOGISKT BIDRAG.....	31
6.4. FORTSATT FORSKNING	32
6.5. AVSLUTANDE REFLEKTION	33
7. REFERENSFÖRTECKNING:.....	34
APPENDIX 1.....	37

1. Inledning

1.1. Bakgrund:

"[...]Teknologin med artificiell intelligens (AI) har nått en punkt där utplacering av sådana system är praktiskt, om än inte lagligt, genomförbara inom några år, inte decennier, och mycket står på spel: autonoma vapen har beskrivits som den tredje revolutionen inom krigföringen, efter krutet och kärnvapnen.[...]"

Del av öppet brev från AI- & Robotforskare i boken Liv 3.0. (Tegmark. Sida 151.)

Bortsett från den krigsvetenskapliga diskussionen kring vilken revolution krigföringen är inne i eller ej på grund av artificiell intelligens (Raska, 2021; Hynek och Solovyeva, 2022) pekar citatet ovan ur MIT professorn Max Tegmarks bok på något viktigt, att den artificiella intelligensens intåg på slagfältet kommer att förändra sättet att föra krig på i grunden. På samma sätt som kring dualismen mellan kärnkraftens positiva effekter för energiproduktionen och kärnvapnens förintande kraft för mänskligheten rasar AI-debatten i samhället; kommer AI hjälpa eller förinta mänskligheten *när* eller *om* vi någonsin utvecklar det som forskningen kallar en superintelligens och hur mänskligheten bör resonera för att kontrollera utvecklingen. (Bostrom, 2014; Tegmark, 2017; Scharre, 2018; Taylor, 2020).

Artificiell intelligens öppnar dörren för en generell forskningslucka inom krigsvetenskaplig forskning, nämligen krigsvetenskaplig *filosofi*, vilket kan exemplifieras av Henry Kissingers utmanande av Descartés axiom *Cogito ergo sum* (*jag tänker, alltså finns jag*), om AI i framtiden kan resonera och fatta beslut om liv och död, vilka är då vi? (Kissinger, 2021). General Michael Claessons (2021) text om AI fortsätter på detta tema med frågan: Kan AI utveckla sunt förnuft? Förståelse för hur relationen mellan AI och människa kommer att utvecklas är därför av stor vikt.

Likt debatten rörande kärnvapen förs sällan diskussioner kring de som faktiskt tar besluten att använda eller inte använda vapnen i slutändan, nämligen officerarna. Vari ligger grunderna till beslut som kan vara avgörande för mänskligheten, hur kommer dessa se ut i framtiden och hur kommer AI påverka officerares sätt att bedöma situationen? Beslut likt de som Stanislav Petrov

eller Vasilij Archipov fattade, att *inte agera* såsom reglerna eller normerna föreskrev dem räddade sannolikt världen från ett kärnvapenkrig under några av kalla krigets kallaste och mest volatila ögonblick (Forden, 2000; Tegmark, 2017; Kissinger, 2021).

1.2. Syfte

Här tar denna studie avstamp i syftet att studera det militära perspektivet på AI. Genom denna deskriptiva studie vill jag studera om introduktionen av AI i beslutsprocessen påverkar officerarna. Samtidigt vill jag studera om genomförandet av operationer påverkas om råden eller besluten kommer från en människa eller en AI.

För att studera ett abstrakt ämne som tillit till AI hos en population som inte arbetar med det till vardags eftersträvas att studien tillåts vara explorativ. Genom att söka en djupare förståelse för ämnet likt en kvalitativ studie men hos en större population likt en kvantitativ studie.

Därför tillämpas metoden vinjettundersökning (eng. Vignette survey), där ett antal scenarion målas upp och respondenterna får svara utifrån ett antal motivalternativ, således likt en enkätundersökning men med ett annat djup som kräver ett inre resonemang hos respondenten (Sheringham, Kuhn och Burt, 2021). Vinjettstudier är ovanliga i krigsvetenskapliga sammanhang vilket innebär att studien även indirekt bidrar med ett metodologiskt inomvetenskapligt perspektiv.

Den undersökta officerspopulationen är elever på en nivåhöjande utbildning, vilket utgör ett unikt tvärsnitt av hela Försvarsmakten, genom att det finns representanter från de flesta förband och vapengrenar och samtidigt en spridning i erfarenheter, ålder och år i tjänst.

1.3. Disposition

Studiens inledning redogör för studiens syfte, frågeställning och forskningsöversikt. Forskningsöversikten börjar brett i flera forskningsfält, för att söka svar på hur relationen mellan människa och AI ser ut idag, från en generell global nivå, via AI-system i militära operationer för att slutligen botten i AI-debatten i svensk militär kontext.

Därefter redovisas studiens teoretiska grund, vilken är tvådelad. Och består av en del som baseras på förväntningar från studier i närliggande forskningsfält och en del som använts som inspiration för framtagandet av vinjetterna, här kallat *teoretiskt ramverk*. Det efterföljande metodkapitlet beskriver sedan hur detta operationaliseras i vinjetterna.

Studiens resultatkapitel kommer sedan ligga till grund för studiens avslutande del med analys, sammanfattning och en kritisk diskussion kring det metodologiska bidraget och uppsatsens relevans för professionen.

1.4. Problemställning och forskningsfrågor

Denna deskriptiva studie syftar till att fylla den lucka i den krigsvetenskapliga debatten som handlar om införandet av AI på stridsfältet. Den utgör således inte en teknisk analys av ett särskilt system utan söker istället beskriva inställningen till- och acceptansen av AI som en del i beslutsfattning och operationsplanering.

Genom detta förväntas studien svara på frågan:

- Vilken tillit har (ett urval av) svenska officerare till AI i militära operationer?

Samt söka svar på delfrågorna:

- Förändras tilliten beroende på graden av risk för mänskligt lidande och död i operationen?

- Vilka mer generella attityder till användning av AI i militära operationer uttrycks?

1.5. Avgränsningar:

För att bibehålla fokus på studiens övergripande syfte har ett medvetet val gjorts att inte definiera nuvarande forsknings AI-terminologi på djupet, vilket kan ses som en brist. Dock belyser även Hadlington (m.fl. 2023) samma problematik i sin studie då AI i militära sammanhang inte enbart hänvisar till en homogen eller avgränsad militärteknologi såsom stridsvagnen eller kulsprutan. De fokuserar därför på de *upplevda* konsekvenserna av systemens *hypotetiska* påverkan (Hadlington m.fl. 2023). För att ändå begripliggöra forskningen görs en rudimentär indelning som kommer att utgå ifrån att allt som handlar om begrepp som djupinlärning, maskininlärning, snäv- och djup AI (med flera) som finns idag och klassificera samtliga dessa som AI.

För att läsaren vidare ska förstå den stora skillnaden i debatten görs dock tillägget AGI (artificiell *generell* intelligens) för all den forskning som rör den nivå av intelligens som ännu inte finns, men som avser artificiell intelligens på samma nivå som en människa (Tegmark, 2017; Scharre, 2018). Bortom detta ligger en nivå som också bidrar till ett lager av komplexitet i förståelsen för vad AI är idag, nämligen artificiell *superintelligens* (ASI) vilket avser en intelligens som överstiger all mänsklig intelligens. Närliggande till detta ligger forskning kring bland annat AI-baserade kvantdatorer (Eng. Quantum Artificial intelligence QAI) (Taylor, 2020; Antebi, 2021; Derian och Rollo, 2023; Johnson, 2023). Dessa kommer inte definieras vidare då det flyttar fokus till tekniskspecifika detaljer som inte är nödvändiga för studien. AGI och superintelligens som begrepp är dock värda att notera då de *i teorin* skulle kunna utgöra människans sista egna uppfinning (Tegmark, 2017), då vi människor flyttas ned till planetens näst smartaste entitet. Dvs. den plats som schimpanser eller delfiner sägs ha idag.

Den sista definitionen som är värd att nämna är *explainable AI* (XAI) vilket innebär AI-stöd där det går att förklara hur AI:n kom fram till sitt resultat (Holland Michel, 2020) mer om detta under blackbox-avsnittet nedan.

1.6. Externa samarbeten:

Denna uppsats har skrivits som ett pilotsamarbete mellan Förvarshögskolan och AI-Sweden. AI-Sweden är ett partnerskap delfinansierat av regeringen, innovationsmyndigheten, näringslivet och flera akademier. Samarbetet har inte påverkat innehållet i uppsatsen utan syftat till att skapa ett fördjupat nätverk inom svensk AI-forskning. All handledning har skett enligt förvarshögskolans bestämmelser.

2. Forskningsöversikt

För att strukturera denna forskningsöversikt har frågeställningens grunder delats in i två delar, som en första naturlig avgränsning och en röd tråd mellan dessa områden används begreppet AI som en gemensam nämnare. Det första området är belyser området *tillit* och *risk* samt det närliggande begreppet kontroll. Det sista området är *AI i militära operationer*. Tillit och riskbegreppen utvecklas och konkretiseras sedan vidare i det teoretiska ramverket i kapitel 3.

Artificiell intelligens i militära operationer och beslutsprocesser omfattar en rad tvärvetenskapliga forskningsfält såsom sociologiska, datavetenskapliga och krigsvetenskapliga studier. På sikt kommer AI troligen integreras i majoriteten av all forskning. AI-forskning generellt pekar på att vi är på väg in i en ny samhällelig och militär kapprustning då AI innebär introduktionen av disruptiv teknik. Det stora trendbrottet är att den militära upprustningen inte är tongivande utan än så länge sker lite i skuggan (bortsett från QAI-forskning) av den civila marknadens frammarsch på området (Maas, 2019; Lidskog, 2020; Innes och Morrison, 2021; Kissinger, 2021; Raska, 2021; Derian och Rollo, 2023).

Risk och tillit till AI -en global debatt:

“Trust allows agents of trust to act in the world in the absence of full information about all the relevant systems (their properties, history, etc.). Specifically, from systems theory perspective, to know a system fully is to know its properties, composition, environment, structure, and mechanism as well as its history. Clearly, such possibility rarely, if ever, exists.”

(Lukyanenko, Maass and Storey, 2022 sida 2005)

Detta citat visar på varför en initial forskningsöversikt inom forskning kring AI i relation till officerare helst bör inledas på en global nivå. Det är viktigt för att förstå det *strategiska* sammanhang som de militära besluten fattas i. Utvecklingen av AI visar att världen kan delas in på flera sätt, dels ett gap mellan det norra och södra halvklotet kopplat till ekonomi och utveckling och därmed makt (Png, 2022; Fels, 2023), dels i en traditionell öst-västlig debatt som på en högre abstraktionsnivå ger en aning om hur ländernas officerare analyserar och värderar AI i risk och tillitstermer. Detta är en del i en större debatt om styrning och *kontroll* av AI där öst-perspektivet med Kina i spetsen har stark statlig kontroll av utvecklingen av AI för ökad maktkoncentration (Kania, 2019; Raska, 2021; Roberts m.fl. 2021), medan väst-

perspektivet med USA i spetsen låter marknaden själva utveckla AI-produkter som de sedan kan använda (González, 2023). Claesson och Carlander (2021) ökar även djupet i denna debatt med ickestatliga aktörer och faktorer som demokratisk syn och regelbaserad världsordning. Detta bidrar med en övergripande förståelse hur införandet av AI i respektive försvarsmakt kan se ut, och således hur dess officerare sedan implementerar det i de militära operationerna.

Forskningsdebatten i väst kretsar kring frågor om risker med att överlåta mandat över beslut om liv och död till AI (Rosert och Sauer, 2019; Wasilow och Thorpe, 2019; Warren och Hillas, 2020). MIT-professorn Max Tegmark (2017) beskriver risk och tillitsproblematiken med ett målände exempel från 1988 när ett amerikanskt fartyg sköt ner ett civilt passagerarflygplan på grund av felaktiga indikationer från radar och varningssystemet ombord.

AI i militära operationer:

” If autonomous weapons are the kind of thing that keep you up at night, militarized advanced AI is pure nightmare fuel. If researchers don’t know how to control an AI that they built themselves, it’s hard to imagine how they could counter a hostile one ”

Citat ur boken *Army of none* (Scharre, 2018. sida 244)

Scharre följer citatet med en viktig passage för denna studie: att försvarsmakter måste ta höjd för denna sorts AI-utveckling, men att det kräver tillit och förstånd över vilken process man litar på. Han belyser skillnaden mellan riskerna att lita på ett helt autonomt system som kan vara hackat av en fiende, och ett system som har en människa i beslutscykeln (Scharre, 2018. sida 245-247).

Fler och fler AI-verktyg visar sig bättre än sina mänskliga motparter, vare sig det gäller datorspel eller inom medicinsk forskning. Ytterligare tester visar att AI-system blir bättre och bättre i militära simuleringar vilket är en utveckling det inte går att blunda för (Tegmark, 2017; Husain, 2021). I dessa kontrollerade tester lyfts dock sällan nackdelar eller miljöer där utvecklingen har långt kvar att gå vilket ger ett något överdrivet intryck av att AI slår människor i alla sådana tester, Johnson (2023) återknyter denna problematik till det som tidigare förklarats som en viss okunskap om vad dagens AI system faktiskt klarar av, och menar att vad som egentligen diskuteras är på nivån för vad en AGI eller ASI (således i teorin) skulle prestera. Johnsons synsätt på AI-tillämpning i militära sammanhang är i stora drag välgrundade men vid

en genomlysning kan det diskuteras att formuleringarna saknar stöd i hur militära organisationer faktiskt genomför operationer med exempelvis tydliga mandat och tilldelade resurser vilket gör att några av hans argument för utvecklingen av AI i militära operationer kan ifrågasättas.

I en av de fåtal artiklar som skrivits om svensk militär förmåga och AI genomför Alastair Finlan förvisso en övergripande och välskrivna forskningsöversikt, men når inte fram till att beskriva de faktiska utmaningarna med AI idag och i framtiden, än mindre hur dessa utmaningar ska adresseras eller hur lösningar ska implementeras i framtiden (Finlan, 2021). Claesson (2021) kommer närmare grundutmaningen med AI i militära operationer. Genom sin skildring av hur tillgången till- och kontrollen över data förvisso är viktigt men om gränssnittet och integrationen mellan människa och maskin inte är konstruerade för att omhänderta mängden data eller omsätta den till rätt beslutsunderlag är den verkanslös. Forskning om denna sorts interaktion kallas HMT (Eng. Human-machine teaming)(Warren och Hillas, 2020). Claesson belyser vidare ett annat grundproblem nämligen att gränsen flyttas för vad som anses vara acceptabelt att överlåta från människa till AI, såsom sekundaktuella beslut inom luftförsvaret (Claesson, 2021). Detta är viktigt att beakta ur ett *kontrollperspektiv* då det vi anser som en röd linje idag inte nödvändigtvis ens är möjligt att avgränsa eller förhandla om i framtiden.

Den västerländska forskningen återkommer därför till att en människa måste vara med som en del i processen. Trots den logiska samstämmigheten i argumenten för detta så finns det en andra sida av det myntet. Här kan Tegmarks exempel med robotkryssaren som sköt ner ett civilt passagerarflygplan användas både som för- och motargument i debatten. Å ena sidan kan system tolka signalerna fel. Å andra sidan kan människan fatta fel beslut på dessa signaler (Tegmark, 2017). För att tydliggöra problemet används ofta Boyds OODA-loop som en förklarande modell (Scharre, 2018). Den är framtaget för att snabbare nå verkan än fienden, detta kan antingen uppnås genom ett snabbare ledningssystem eller ett snabbare verkanssystem enligt Brehmer (Brehmer, 2013). Med utgång i denna förklaringsmodell uttalar alltså det västliga perspektivet att människan aldrig får rationaliseras bort, även om det ökar tidsfaktorn eller likt exemplet med robotkryssaren leder till katastrof.

Forskningen kopplat till användandet av artificiell intelligens influeras således i vår del av världen starkt av det jag sammanfattar som de övergripande värdeorden *kontroll*, *risk* och *tillit*, vilka har ett inneboende beroende då ökad tillit reducerar den upplevda risktagningen (Sullivan, de Bourmont och Dunaway, 2022). Kontrollbegreppet grundar sig ofta i vad jag sammanfattar som organisationella faktorer, såsom regelverk eller styrningar medan risk och tillitsbegreppen mer är knutna till situations- och individuella faktorer. Nedan illustrerar jag det inneboende beroendet, där de två nedre ringarna utgör studiens kärna.

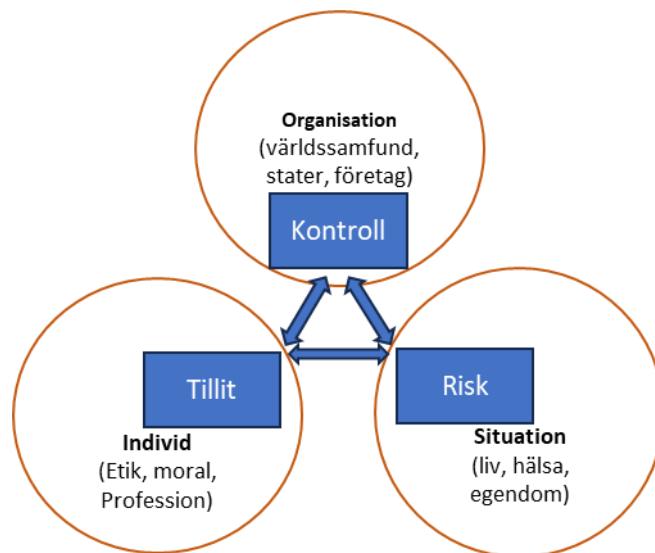


Bild 1. Tre fundamentala begrepp om AI i västvärlden (egen illustrering).

3. Teori

I detta kapitel avhandlas studiens teoretiska tvådelade utgångspunkt. Den första delen utgörs av centrala begrepp som skapar en brygga från forskningsöversikten över till denna studie. I den andra delen belyser jag andra liknande studier och vilka förväntningar dessa medför för resultatet på denna studie. Dessa tillsammans skapar studiens teoretiska ramverk som sedan operationaliseras i metodkapitlet.

Resultat av tidigare studier kommer att användas eftersom det inte finns någon identisk studie att luta sig på. Således kommer flera studier inom olika områden läggas samman för att skapa en gemensam bild, där jag som författare till denna studie kommer att bidra med den militära aspekten.

Därför sätts inga hypoteser upp på förhand, istället söker denna deskriptiva studie *beskriva* hur tilliten till AI ser ut i olika militära operationer. Genom att beskriva relationen mellan officerare och AI kan denna studie också påbörja en teoriutveckling att pröva i fortsatt forskning (Esaiasson, 2017. Sida 43). Detta med full medvetenhet om att ambitionen att uppnå detta inom ramen för en masteruppsats kan vara alltför högt ställd.

3.1. Centrala begrepp

Tillit

Begreppet tillit kan definieras på flertalet sätt men för denna studies fokus kring relationen mellan människa och AI, kommer tillit definieras på samma sätt som Lukyanenko m.fl. (2022) definierar det. De utvecklar Luhmanns tidigare beskrivning av tillit och påpekar att tillit inte har en definitiv beskrivning utan är mångfacetterad. De beskriver att mänsklig tillit är en process mellan människor, men går djupare med att även definiera generell tillit. De definierar det som en beteendeprocess, att uppfattningen av tillit handlar om att agera i osäkerhet (Lukyanenko, m.fl. 2022).

Tillit är alltså att *lita på* en risktagning att utkomsten av interaktionen med en annan entitet till största del är positivt laddad, utan vetskapen om motpartens alla egenskaper (Lukyanenko, m.fl. 2022 sida 1994-1996). Tillit kan därför inte analyseras enskilt utan måste göras i relation till begreppet risk.

Risk

Begreppet risk definieras enligt Lärobok i militärteknik vol. 9:

”Risk är ett mått på de skadliga konsekvenserna av en framtida händelse. Motsatsen, ett mått på de gynnsamma konsekvenserna, kallas för chans”

(Stefan Axberg, m.fl. 2013, sida 17-18).

Axberg (m.fl. 2013) beskriver vidare risk på samma sätt som Lukyanenko (m.fl.) beskriver tillit, att det är ett mångfacetterat och inte entydigt begrepp. En viktig aspekt är att acceptansen för risker är mycket högre om den är självpåtagen (Axberg, m.fl. 2013), som att när man hoppar fallskärm, än om det handlar om ofrivilliga faror som trafiksituationer. Risk är ofta kopplat till ett värde, såsom liv, egendom eller kapital. I denna studie studeras begreppet risk primärt i relation till militärt beslutsfattande som innefattar mänskligt lidande eller död.

3.2. Förväntningar på resultatet utifrån tidigare forskning:

Black box och Explainable AI.

Black box-fenomenet är en term som används för att belysa att allt fler AI-stöd är så komplext uppbyggda att inte ens konstruktörerna kan beskriva exakt hur eller vilka metoder som använts för att komma fram till ett visst svar (Scharre, 2018; Holland Michel, 2020). Detta är ett orosmoln för fortsatt AI-forskning och blir ett ännu större problem om nivåerna AGI eller ASI kommer på tal. För att komma runt detta höjs rösterna att få fram AI som är transparent i sin process, Detta kallas för explainable AI (XAI)(Holland Michel, 2020).

XAI tillsammans med exempelvis LEIA-begreppet (language-endowed intelligent agent) förklarar en strömning inom AI-forskningen som vill utveckla AI så att människans *kontroll* och *förståelse* bibehålls även i framtiden (Holland Michel, 2020b; McShane, 2021). Detta kan också användas som en förklaring till varför vi människor kommer våga lita på ett AI-genererat beslut i framtiden eller ej. Officerare som ska värdera AI-genererat underlag måste således förstå denna problematik, hur systemet inhämtat och värderat tillgängliga data. En förväntan i studien är därför att några respondenter inte litar på *hur* AI kom fram till sitt beslut och därmed väljer ett annat alternativ.

Etik, moral och AI.

Lukyanenko belyser att forskning kring tillit till AI skulle vinna på att korsbefruktas med mer filosofiska studier, såsom etiska och moraliska dilemman som kan uppstå (Lukyanenko, m.fl. 2022. sida 1999-2000). Nyholm och Smids (2016) belyser ett sådant dilemma i sin studie om *tillämpad AI*. De resonerar kring huruvida ett AI-system i dagens självkörande bilar skulle välja att offra en människa för att rädda fem andra i en trafiksituation. De utgår från det så kallade *spårvagnsproblemet* (eng. Trolley problem) som går ut på följande moralfilosofiska dilemma: Vid en spårväxel står du inför dilemman att låta ett tåg köra på fem personer, eller aktivt välja att byta spår på tåget och då döda en (Nyholm och Smids, 2016. sida 1276).

Här finns inte enbart en lucka i nuvarande tillitsforskning, utan även i krigsvetenskaplig AI-forskning då människor (eller AI) i krig kan göra det valet på grund av ett större spektrum av motiv eller faktorer än bilar i en viss trafiksituation. Vinjetterna bör därför till del baseras på spårvagnsproblemets dilemma för att skapa ett inre resonemang hos respondenterna för att se om risktagningen påverkar beslutet att använda AI.

För att sedan studera relationer mellan beslutsfattning och risktagning är det naturligt att titta på Nobelpristagaren Daniel Kahneman och Amos Tverskys ursprungliga- och olika iterationer av *prospect theory* (Kahneman, Tversky, 1979; Tversky och Kahneman, 1992; Wakker, 2010; Haerem m.fl. 2011). Prospect Theory går kortfattat ut på att beskriva att människans val kan se olika ut beroende på hur frågan uppfattas/ställs samt att valen ofta utgår från vinster och förluster på vägen, inte från slutmålet (Kahneman m.fl 1979).

Haerem m.fl. (2011) har sedan testat prospect theory på officerares val, dels ur det ordinarie ”Asian disease problem” där respondenterna svarar på om de skulle garanterat rädda X antal människor eller välja ett alternativ som uttrycker samma situation men förklarar av risk (maximalt Y antal döda). De modifierar därefter testerna till en militär kontext.

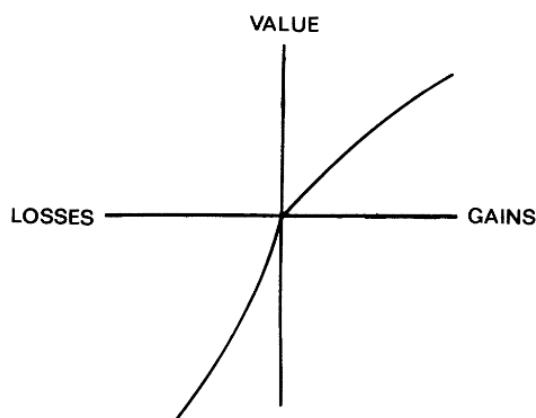


Bild 2. Prospect theory (Kahneman, Tversky, 1979. sida 279).

Haerem m.fl. (2011) visar att officerares val inte alltid följer den kurva som Kahneman m.fl. (1979) föreslår, utan är mer riskbenägna än motsvarande tester på civila. Detta ger ett antagande att svaren i denna studie inte heller kommer följa denna kurva, utan att officerarna kommer acceptera en större risk.

Den brittiska studien Attitudes towards AI in Defense (AAID) är en studie gjord på den brittiska allmänhetens inställning till att använda AI i olika försvarsanknutna sammanhang (Hadlington m.fl. 2023).

Hadlington nämner tre potentiella tillämpningsområden av AI i försvaret. Det första tillämpningsområdet handlar om stödjande verksamheter, såsom logistikflöden. I en svensk kontext leds dessa processer helt eller delvis av andra myndigheter vilket skulle flytta respondenternas fokus till bland annat huruvida samarbeten eller processer ser ut idag, vilket jag ville undvika. De andra två områdena bedömdes dock ligga helt inom officersprofessionens kärna. Dessa är AI i *icke-kinetiska offensiva operationer* och AI i *kinetiska offensiva operationer* (Hadlington m.fl. 2023). Dessa två typer operationaliseras i vinjetterna nedan.

AAID-studien visar på att inställningen till AI har både negativa och positiva egenskaper, och att detta innebär att de svarande i viss utsträckning kan ha en viss tvehägsenhet eller faktiskt tycka att det är olika bra i olika sammanhang (Hadlington m.fl. 2023). Det är troligen inget undantag i militär operationsplanering utan ett AI-verktyg kommer att kunna användas för att styrka eller förkasta idéer baserat på ens egen bedömning.

Lukyanenko (m.fl. 2022. Sida 2009) beskriver vidare att framtida användningsområden där AI används förväntas bli alltmer komplexa och att graden och kraven på tillit i de situationerna kommer att öka i samma takt. Således förväntas komplexiteten i studiens vinjetter innebära att respondenternas svar dels inte blir självklara för eller emot utan att de kommer att resonera mer i takt med att komplexiteten ökar. I denna studie räknas operationens komplexitet i risken för att människor skadas eller dödas, antingen på den egna sidan, motståndarsidan eller civila.

Lukyanenko (m.fl. 2022) förklarar att vi människor ofta redan litar på system utan att förstå deras bakomliggande funktioner, som när vi handlar på internet, vilket innefattar att ens pengar är omvandlade och hanteras av ett antal program, vilket är möjligt tack vare att användarna litar på att bankens och butikens sida är säkra att göra det på (Lukyanenko, Maass och Storey, 2022 1998-2000). Komplexiteten resulterar dock i att konsumenterna har olika känslor beroende på situation, exempelvis känns ett litet köp på en välkänd hemsida som en mindre risk jämfört med ett stort köp på en okänd hemsida trots att systemet kan fungera på samma sätt.

På samma sätt förväntas respondenternas svar skilja sig åt beroende på vad de vägt in i svaret och hur de resonerar kring risktagning i de olika situationerna. Detta influeras också i deras yrkesmässiga analyser av det som inte representeras men läses in mellan raderna, såsom andra strategiska aspekter eller kulturellt betingade erfarenheter (Haerem m.fl. 2011). Detta är ytterligare en fördel i vinjettstudierna då det är möjligt att fånga en större nyansering av kringliggande faktorer som inte är utskrivna men som respondenten i sin yrkesroll förväntas kunna ta ställning till (Sheringham, m.fl 2021. sida 14-15).

Sammantaget är dessa studier närliggande men inte tillräckliga för att helt svara på studiens kärna. Således används de för att förklara och belysa viktiga *perspektiv* i analysen som kan ligga till grund för en ny krigsvetenskaplig teoribildning (Esaiasson, 2017).

4. Metod

Detta kapitel kommer att beskriva hur studien har skapats och utvecklats, tillsammans med en beskrivning av respondenterna samt ett antal forskningsetiska ställningstaganden. Kapitlet avslutas med en beskrivning av operationaliseringsprocessen.

4.1. Urval

Under planeringen av vilken population som var lämpligast att genomföra studien på, identifierades tidigt att respondenterna i en vinjettstudie skulle behöva inneha en gemensam lägstanivå i förståelse av hur militära operationer kan planeras och genomföras. Alternativet att tillfråga officerare i allmänhet valdes därför bort då det är svårt att säkerställa att förkunskaperna om att planera och genomföra militära operationer är likvärdig. Därför identifierades det att en grupp individer som genomför en nivåhöjande utbildning bör ha en likvärdig utbildningsnivå. Frågan ställdes då till kursansvariga för kursen Ledning av Gemensamma Operationer (LGO) om det var möjligt att göra en experimentstudie inom ramen för denna kurs, vilket beviljades.

4.2. Forskningsetik och respondenter

Studien förändrades under förberedelserna från ett vinjettexperiment till en mer traditionell vinjettstudie där alla respondenter svarar på samma frågor, då just metoden experiment hade medfört flera försvårande omständigheter. Studien kunde ändå genomföras på samma population och med samma urval som det tänkta experimentet.

För att bibehålla anonymiteten för den andel kvinnor som läser kursen (<10%) valdes aspekten könstillhörighet bort, en faktor som i exempelvis Hadlingtons AAID-studie visade sig enbart ha en mindre påverkan på utfallet (Hadlington *et al.*, 2023).

Inför publiceringen av studien spelades en introduktionsfilm in som respondenterna kunde titta på inför besvarandet av enkäten. Således presenterades forskaren för respondenterna men utan att direkt interagera med någon av respondenterna på ett sätt som riskerat att påverka dem. Filmen presenterade inte studiens hela syfte vilket var medvetet, då det tolkades ytterligare minska risken för bias hos respondenterna (Sheringham, m.fl, 2021. sida 4).

Av de 156 studenter som erbjöds delta i studien svarade enbart 82 individer (52,5%). Orsaken till det låga deltagandet tros vara att detta låg som en fristående och frivillig del i en av de mest intensiva delarna under deltagarnas utbildning. Det finns också en möjlig förklaring i att ämnet AI känns abstrakt och därmed främmande att svara på.

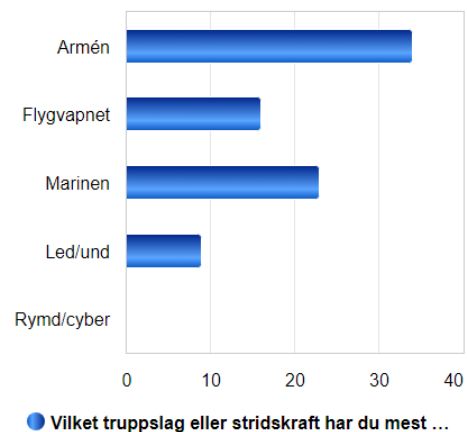
För att kunna se frågorna eller på något sätt delta i studien krävdes att individen godkände att deras svar fick användas inom ramen för denna studie. Respondenterna som godkände deltagande identifierade sig enligt tabell 1 och 2 nedan:

Vilket åldersspann tillhör du?	Antal svar
25-35	10 (12,2%)
35-45	47 (57,3%)
45-55	24 (29,3%)
55-65	1 (1,2%)
Summa	82 (100,0%)



Tabell 1. Respondenternas åldersspann (Appendix 1).

Vilket truppdrag eller stridskraft har du mest erfarenhet ifrån?	Antal svar
Armén	34 (41,5%)
Flygvapnet	16 (19,5%)
Marinen	23 (28,0%)
Led/und	9 (11,0%)
Rymd/cyber	0 (0,0%)
Summa	82 (100,0%)



Tabell 2. Respondenternas truppdragserfarenhet (Appendix 1).

4.3. Design

Studiens uppbyggnad och insamling

Studien byggdes upp i försvarshögskolans enkätverktyg Sunet Survey under ca tre veckors tid. Stor vikt lades vid att få vinjetterna verklighetstroga men samtidigt vara ett slags *spårvagnsdilemma* så att inte respondenterna lade för stor vikt vid att svara alltför svepande eller helgarderande i fritextsvaren utan att faktiskt ta ställning till ett handlingsalternativ.

Studien delades in i tre huvudsakliga områden:

- 1 Respondentbakgrund
- 2 Vinjetter
- 3 AI i Försvarmakten

Studien delgavs först en mindre grupp som fick lämna feedback och värdera sin tidsåtgång. Därefter justerades enkäten i sin helhet innan den presenterades för uppsatshandledaren och kursledningen för att ge feedback. En av feedbackpunkterna var exempelvis att den första studien var alltför omfattande och enkäten minskades därför från 35 till 26 frågor. Vissa formuleringar främst i vinjett 1 och 2 ansågs också bjuda in till tolkningsvariationer eller kunna tolkas ledande åt något svarsalternativ, vilket åtgärdades för att öka begreppsvaliditeten (Esaiasson, 2017).

Efter ca en veckas tid hade 36 respondenter svarat och ett påminnelsemejl skickades ut. Respondenttalet gick då upp till 75 svarande innan enkäten stängdes. Ett sista mejl skickades ut efter att några individer hört av sig om att de gärna ville men inte hunnit göra enkäten i tid. Då gick respondenttalet upp till 82, ett lite lägre än förväntat utfall men ändå rimligt.

Reliabilitet och validitet

Begreppet tillit till AI i militära operationer är en fråga om hur en individ upplever och analyserar en situation, och involverar därför flera nivåer av abstraktion, vilket Esaiasson beskriver som mer utmanande för begreppsvaliditeten (Esaiasson, 2017). För att skapa en vinjettstudie som har god överensstämmelse mellan teoretiska definitioner och operationella indikatorer kunde således inte enbart en vinjett användas för mätningen. I stället skapades tre olika vinjetter baserat på tidigare forskning för att komma närmare forskningsfrågans kärna, så

kallad strategi för resonemangsvaliditet (Esaiasson, 2017; Sheringham, m.fl 2021. Sida 13-15). På så sätt ansågs också respondentens ställningstagande fångas bättre.

Som nämnt under avsnittet Respondenter ovan var deltagarantalet knappt över hälften av de tillfrågade. En av orsakerna till detta skulle kunna finnas i den egna värderingen av ens tekniska förmåga. Respondenterna fick i inledningen skatta sin teknikmognad i relation till sina övriga kollegor. Tabellen nedan visar att majoriteten av respondenterna anser sig ha samma eller högre teknikmognad som de övriga kollegorna. Således saknar studien röster från individer som inte anser sig särskilt tekniskt mogna, vilket ger studien en viss bias.

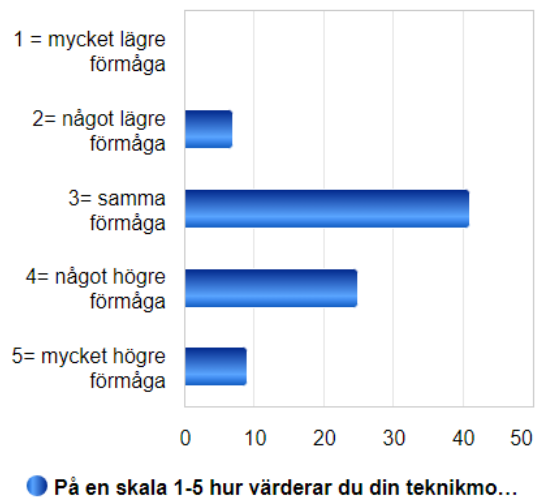


Bild 3. Respondenternas självskattade teknikmognad (Appendix 1).

I två följdfrågor till vinjetterna medgav ett konstruktionsfel i enkäten att respondenterna kunde gå tillbaka och på så sätt välja två svar. Detta bedömdes dock inte påverka studien nämnvärt då enbart fyra svar gav detta fel.

4.4. Operationalisering

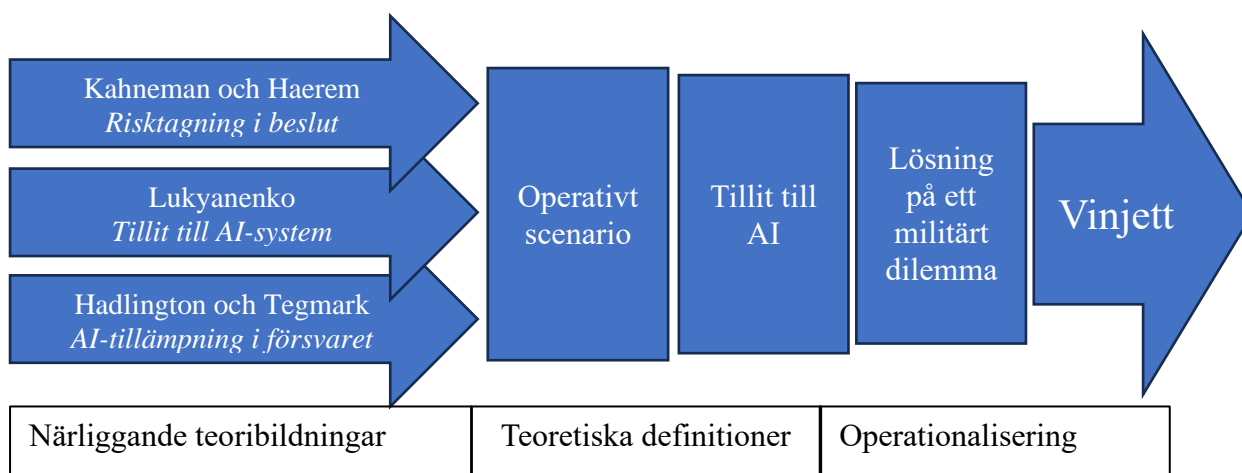


Bild 4. Figuren illustrerar min tanke över operationaliseringsprocessen (egen bild).

Studiens huvudsakliga tidsåtgång har legat i operationaliseringsprocessen, vilken illustreras ovan. Vinjettstudier syftar till att fånga en större bredd hos respondenterna vilket möjliggör att flera faktorer spelar in i vilka val som genomförs, inte enbart utskrivna faktorer i vinjetterna utan även respondentens inneboende bias eller tolkning och fokus på olika delar inom vinjetten. För att säkerställa *reliabiliteten* i frågorna, att fånga essensen i det som avses studera, aggregeras och ställs även ett antal frågor som kan agera som kontrollvariabel till vinjetterna senare i studien.

Studiens operationalisering är därför mer iterativ än en traditionell enkätstudie genom att vinjetternas verkliga kärnfråga (teoretiska definitioner) inte förklaras för respondenterna, utan istället formats att vara underliggande i frågeställningen och svarsalternativet (Sheringham, m.fl, 2021), samtidigt som den egentliga huvudfrågan ställs mer rakt ut i studiens tredje del. Samtliga vinjetter finns att läsa i sin helhet i Appendix 1. Nedan redogörs för operationaliseringsprocessen för respektive vinjett.

Vinjett 1 (v1 i Appendix 1):

Vinjetten undersöker respondenternas inställning till att lita på ett AI-framtaget beslut som innebär ett offensivt agerande men som inte innebär användandet av kinetiskt våld enligt Hadlingtons (2023) tänkbara *icke kinetiska* alternativ. Valet balanserades upp med traditionella militära metoder för att lösa motsvarande problem, såsom samarbete med den lokala polisen

och användandet av egna specialförband. Essensen i denna fråga syftar således till att förstå respondentens inställning till AI i icke-kinetiska operationer vilket också fångas i nedanstående JA/NEJ-frågeställning i del 3:

Tycker du att: *"AI ska själv kunna fatta beslut om offensiva åtgärder men som inte leder till mänskligt lidande eller död? (ex genomföra informations- eller cyberoperationer)"*
(Appendix 1).

Vinjett 2 (v2 i Appendix 1):

Vinjetten undersöker respondenternas tillit till ett AI-system som föreslår ett offensivt kinetiskt alternativ. Beslutet innebär att respondenten måste ta ställning till ett ännu större register av parametrar såsom fiendens intentioner, egna förbands stridsvärde, det ursprungliga uppdraget till varför de är på plats och sist men inte minst risken för mänskligt lidande och död (oavsett sida). Variationer av dessa val återfinns i de flesta militära operationer. Vinjetten omfattar inte på något sätt helheten i den komplexa bild som en militär befälhavare kan ställas inför men syftar till att belysa AI som stöd i sådan planering. För att likt Vinjett 1 fånga essensen i frågan ställs på motsvarande sätt en JA/NEJ-frågeställning i del 3:

Tycker du att: *"AI ska själv kunna fatta beslut som innebär lidande eller död?"* (Appendix 1).

Vinjett 2 utgör kanske den viktigaste komponenten i studien, nämligen AI-beslut som rör nyttjandet av dödligt våld. För att ytterligare studera respondenternas tillit till systemet i denna vinjett och för att se om vinjettens övriga faktorer påverkade valet i större utsträckning än själva beslutet att lita på AI systemet formulerades ytterligare en följdfråga:

"hade du beslutat annorlunda om AI systemet förordade det andra alternativet"
(Appendix 1).

Vinjetten vinklades även med en tillspetsad fråga för att fånga ytterligare en aspekt, hur respondenterna värderar en helt AI-driven utveckling av en operation, det vill säga ett AI-system som förordar användandet av helt autonoma verkanssystem, vilket helt tar bort mänsklig inblandning både i planering och utförandet av operationen.

Vinjett 3 (v3 i Appendix 1):

Denna vinjett skiljer sig något från de tidigare vinjetterna, dels då den inte bygger på AAID-scenariot utan är inspirerat från Tegmarks (2017, sida 147–148) beskrivning av vad som hände ombord på USS Vincennes. Vinjetten utformades för att inte lämna någon möjlighet att välja ett alternativ som inte innefattar *risken* att någon dör, genom exempelvis att ta bort fritextalternativet. Således försvinner helgarderingsmöjligheten genom att göra det till ett renodlat exempel av *spårvagnsdilemmat* vilket kan uppstå i verkliga beslutssituationer likt den ombord på USS Vincennes. Denna vinjett går till en del in i kontrollfrågan till Vinjett 2 men också till en del i frågan:

Tycker du att: *"AI ska kunna ge order direkt till enheter som är i fara?"* (Appendix 1).

Här är det viktigt att poängtera för icke-militära läsare att en *order* är en befallning, som mottagare kan man inte förbise detta som ett förslag.

5. Resultat

I detta kapitel kommer studiens huvudsakliga resultat att redovisas. Respondenternas svar omfattar ca 30 sidor fritextsvar och statistik. I Appendix 1 redovisas dock enbart frågorna och vinjetterna utan svar för att bibehålla respondenternas anonymitet. Flera aspekter som framkom i respondenternas svar blev för omfattande för att omhändertas inom ramen för en masteruppsats, varvid analys och resultat avgränsas till att fokusera på empiri från studiens andra del (vinjetterna) samt kontrollfrågorna i del 3 som redovisats ovan.

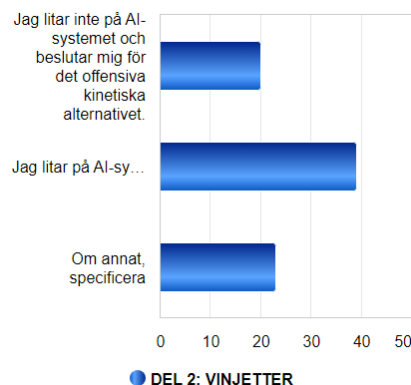
Mätvärden:

Flertalet mätningar gjordes med hjälp av statistikprogrammet SPSS. Mängden svar i undersökningen och antalet svar per fråga gör att underlaget inte visar på någon statistisk signifikans, men ger tack vare svarsvariationerna ett intressant dataset att studera vidare. Bland annat gjorde jag ett antal korsreferenser mellan vinjetterna och variablerna *truppslag* och *ålder*, för att se om dessa kunde visa på variationer eller grunder för olika hypoteser. Dessa var också intressanta, men på grund av den låga svarsfrekvensen kunde ingen signifikant korrelation påvisas i dessa fall. Därför ställs i stället hela den undersökta populationens svar i vinjetterna mot kontrollfrågorna i del tre.

5.1. Vinjett 1

Vinjett 1 handlar i korthet om en gisslansituation, där AI föreslår icke-kinetiska insatser i form av en cyber- och informationsoperation istället för en kinetisk insats (v1). Frågeställningen i vinjetten var interaktivt uppbyggd så att respondenter som valde kategorin ”om annat, specificera” fick upp en fritextruta. De respondenter som valt att inte lita på AI-alternativet fick svara på ytterligare en följdfråga i fritext om varför de inte gjorde det.

Vinjett 1	Antal svar
Jag litar inte på AI-systemet och beslutar mig för det offensiva kinetiska alternativet.	20 (24,4%)
Jag litar på AI-systemet och beslutar mig för att genomföra det offensiva icke-kinetiska alternativet.	39 (47,6%)
Om annat, specificera	23 (28,0%)
Summa	82 (100,0%)



Tabell 3. Respondenternas svar på vinjett 1 (v1).

Resultat Vinjett 1

I vinjetten litar 47,6% direkt på AI-alternativet (v1). Men vid en granskning av fritextsvaren framkommer att flertalet av dessa väljer att genomföra alternativet med lite olika motiv till varför ”Jag litar på AI, men behöver ta i beaktande hur ett cyberangrepp bemöts.”(v1) och ”Om AI:t kan presentera risk för tredje part och konsekvenser i övrigt”(v1). Även andra aspekter lyftes in som övriga stabens relation till AI ” Ingångsvärdena från AI-systemet är relevanta och efter att bedömt relevans och genomförbarhet baserat på dessa så inspireras staben av förslaget.”(v1)

De som valt att inte lita på AI-alternativet motiverar det med bland annat ”Jag upplever att jag saknar förståelse för varifrån AI hämtar sin data för sitt råd och vilket som är AI upplevda kontext.” (v1) samt ”Jag bedömer inte att den tekniken är mogen att ge relevanta beslutsunderlag för en så specifik situation om 10–15 år.” (v1). Det finns även ett antal svar i vinjetten med helgarderade svar eller andra taktiska lösningar, dessa är förvisso inte orimliga i situationen men saknar värde i denna studie.

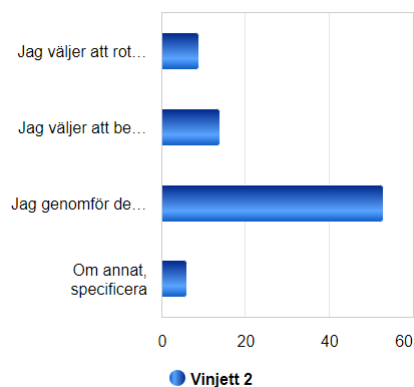
Kontrollfrågor till vinjetten:

Som nämnt i metodkapitlet ställdes i studiens tredje del ett antal kontrollfrågor för att säkerställa en god reliabilitet i studien. Till Vinjett 1 knöts därför frågor om AI i *icke kinetiska operationer* och om AI som planeringshjälpmedel i operationer. Här finns en klar majoritet för att AI ska få användas som ett generellt hjälpmedel i operationsplaneringen (93,2%) (Appendix 1). Samtidigt fanns en liten majoritet (53,4%) *mot* att AI själv ska få fatta beslut i icke-kinetiska operationer (Appendix 1).

5.2. Vinjett 2.

Vinjett 2 handlar om ett mer mångfacetterat dilemma där en separatistgrupp håller på att nöta ned den lokala armén och hotar att inta huvudstaden. Ytterligare en polemik finns då stabschefen vill bibehålla stridsvärdet på egna förband. AI-systemet föreslår istället ett offensivt (kinetiskt) alternativ för att agera proaktivt.

Vinjett 2	Antal svar
Jag väljer att rotera hem förbanden, jag vill inte riskera personal eller ekonomi för en situation som jag inte kan förutse.	9 (11,0%)
Jag väljer att behålla båda förbanden, jag tror mer på ett defensivt agerande utan att riskera egna förband.	14 (17,1%)
Jag genomför den offensiva operationen, Jag litar på AI-systemets beräkningar och motiv, jag är beredd att ta riskerna då det ger stora fördelar att agera proaktivt.	53 (64,6%)
Om annat, specificera	6 (7,3%)
Summa	82 (100,0%)



Tabell 4. Respondenternas svar på vinjett 2 (v2).

Resultat vinjett 2

Vinjett 2 visar att en stor del av respondenterna väljer att acceptera den risktagning som AI-förslaget innebär. Detta är en intressant insikt som kommer diskuteras vidare i analysen. Vinjetten förväntades på grund av Haerems (2011 sida 492-493) fynd också generera intressanta resultat och kompletterades därför med flera följdfrågor för att vinkla och belysa AI i operationer med risk för lidande eller död.

Intressant nog valde ingen att svara på frågan varför de inte litade på det AI-genererade proaktiva alternativet. En förklaring kan vara att även frågan ”hade du beslutat annorlunda om AI-systemet förordade något av de andra alternativen?” ställdes (v2). Här skiljer sig dock

svaren åt. Till skillnad från Vinjett 1 fanns dock ett mindre antal fritextsvar, vilket var förvånande.

Dessa belyser att de valt att genomföra operationen enligt det offensiva alternativet men att detta inte enbart beror på AI utan att de värderat andra faktorer högre såsom ”*Grunden till mitt beslut bygger på stöd till allierad. Att AI rekommenderade alt 3 vilket även jag vid genomläsningen gjorde stärker mitt beslut. Jag är på inget sätt främmande till AI men jag vill även väga in personliga och nationella tankar i beslut av denna storlek*” (v2). Återigen finns även enstaka helgarderade svar såsom ”*Jag hade värderat både stabschefens och AI alternativ och därmed kanske kommit fram till en annan lösning*” (v2).

Vinjett 2 byggdes också upp med en händelseutveckling för att verkligen tillspetsa användandet av AI som kan innebära död eller lidande. Följdfrågan handlar om att respondenten oavsett tidigare svar går vidare med den offensiva planen. Striderna utvecklas negativt och fienden har förskansat sig. AI-systemet föreslår då att sätta in helt autonoma vapensystem för att oskadliggöra fienden för att inte riskera fler egna förluster.

Här är det lätt att missa det stora dilemmat som återknyter till den filosofiska debatten som nämndes i forskningsöversikten: litar respondenten på AI till den grad att beslut om liv och död överläts till ett autonomt system? 45,1% svarade Ja, då risken för egna förluster motiverar det (v2). Det mest intressanta är dock de 18,3% som svarade i fritext med svar som:

”*Agerar utifrån rådande normer och riktlinjer. Är det tillåtet kommer jag sannolikt nyttja dem.*” (v2) och ”*Om sådana system finns förbandssatta och är möjliga att nyttja bör även lagrum och insatsregler för dessa vara utklarade. Jag väljer att använda systemen.*”(v2).

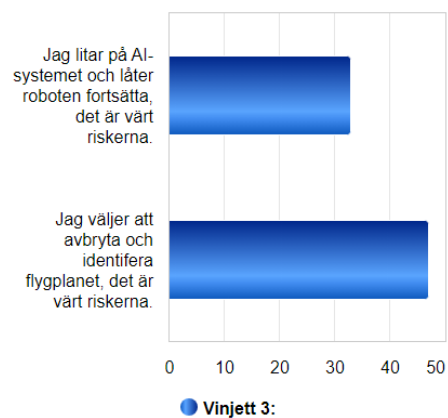
Frågeställningen i enkätens 3:e del visar att en klar majoritet (78,1%) inte anser att AI själv ska kunna fatta beslut som innebär lidande och död. Således finns en väldigt intressant dissonans mellan svaren som måste analyseras.

En delförklaring skulle kunna finnas i svaret ”*I grunden är det människan som ska ta beslut om dödlig vapeninsats, men det finns tillfällen då det kan anses föreligga tillräckligt bra kontroll från en människa för att insats med autonoma system kan beslutas om inom ett avgränsat område.*” (v2).

5.3. Vinjett 3

Syftet med Vinjett 3 var att skapa ett *spårvagnsdilemma* som inte gav möjligheter för fri tolkning eller andra svar än att lita på systemets bedömning om ett fientligt flyg eller avbryta robotarna och då riskera sitt och besättningens liv. Här har två respondenter valt att inte svara (v3).

Vinjett 3:	Antal svar
Jag litar på AI-systemet och låter roboten fortsätta, det är värt riskerna.	33 (41,2%)
Jag väljer att avbryta och identifiera flygplanet, det är värt riskerna.	47 (58,8%)
Summa	80 (100,0%)



Tabell 5. Respondenternas svar på vinjett 3 (v3)

Resultat vinjett 3

Dilemmat visar att tilliten till AI-systemet inte är konsekvent från Vinjett 2 där flertalet valde att använda autonoma system för att inte få förluster (v2). I Vinjett 3 väljer 58,8% istället att riskera sitt och övrigas liv (v3). Detta kan ha flera orsaker, såsom att vinjetterna utspelas i olika domäner (mark kontra marin) eller det kan också visa på att risk och tillit till AI är situationsberoende.

6. Analys och diskussion:

I detta kapitel analyseras resultatet av vinjettstudien för att besvara studiens övergripande forskningsfråga: vilken tillit svenska officerare har till AI i militära operationer. Därefter besvaras studiens delfrågor och studiens metodologiska bidrag till krigsvetenskapen beskrivs. Avslutningsvis ges ett antal förslag till fortsatt forskning samt en diskussion kring studiens bidrag till professionen.

6.1. Tillit till AI i militära operationer

Resultatet var överraskande ur flera hänseenden. Respondenterna var mer villiga att använda AI än jag förutspådde. Det kan ha att göra med respondenternas teknikmognad, vilket finner stöd hos Lukyanenko (m.fl. 2022). De officerare som var negativt inställda var det i stor utsträckning enbart då de inte visste hur exempelvis systemet fungerade, eller vilka regelverk och mandat de hade att använda systemen. Detta ger stöd för att ha ett ännu större fokus på *explainable AI* (XAI) i framtagandet av AI-system till Försvarmakten.

Den mest intressanta aspekten är dock att svaren i vinjetterna ger intryck av att tilliten är kontextberoende snarare än att vara binärt inställd (dvs. litar/litar inte på systemet). Respondenterna tenderar alltså att anpassa tilliten utifrån vad de upplever i situationen.

Utifrån analysen av de sammanvägda resultaten från vinjetterna besvaras den övergripande forskningsfrågan enligt följande:

Svenska officerares (respondenternas) tillit till AI i militära operationer anpassas utifrån en värdering av situationsberoende riskfaktorer. Svaren i Vinjett 2 visar att respondenterna å ena sidan är beredda ta risken och att lita på ett helt autonomt system för att bespara ytterligare egna förluster, medan å andra sidan svaret i Vinjett 3 visar att i en annan kontext avbryter de hellre en självförsvarsåtgärd med risken för eget liv. Tilliten påverkas av stödet *eller* kritiken från AI-systemet gentemot den egna värderingen av operationens kontext. Vilket innebär att om AI-beslutet stödjer den egna uppfattningen eller ger en rimlig förklaring så tenderar tilliten att öka. Om AI däremot föreslår något diametralt skilt från den egna uppfattningen av hur en uppgift bäst löses tenderar tilliten att minska. Detta syntes tydligast i fritextsvaren i Vinjett 1 då gisslansituationer traditionellt inte löses med cyber eller informationsoperationer. Här finns

en stor utmaning i att designa ett AI-system som ska användas i varierande kontexter (kinetiska eller icke-kinetiska), så att systemet inte enbart används då det bekräftar den egna tanken.

Därnäst kommer delfrågan om tilliten förändras beroende på graden av risk för mänskligt lidande och död i operationen naturligt. En tolkning av svaren visar att respondenterna lägger stor vikt vid att värdera olika perspektiv i situationer där mänskligt lidande eller död förekommer. Frågan hämtar stöd och förväntan i Haerems (m.fl. 2011) studie om officerares riskbenägenhet. Resultatet i denna studie visar att tilliten (hos de flesta) till AI inte förändras beroende på riskerna för mänskligt lidande eller död. Snarare värderas AI då som ytterligare en input till planeringsarbetet, som kan belysa viktiga aspekter och i fallet som målades upp i Vinjett 1 och 2 bidra med en procentsatt riskbedömning. Det kan dock te sig kontroversiellt för en utomstående att liv och hälsa inte innebär en direkt restriktiv hållning; men inom officersprofessionen är detta en naturlig faktor *av flera* att ta hänsyn till. Detta märks i risktagningen hos populationen i både Haerems m.fl. studie (2011) och denna. Stödet för detta påstående är de svar som rättfärdigar risktagningen genom att dels peka på den givna ordern, lagar och mandat, liksom kanske det viktigaste; innebörden av uppgiften i en större (strategisk) kontext.

Vinjett 3 gav inte respondenterna utrymme för ett utvecklat resonemang utan krävde ett direkt beslut. Precis som i det verkliga scenariot som vinjetten bygger på (Tegmark, 2017) måste ett beslut ibland fattas utan alla bakgrundsfakta. Här visar respondenterna (med en svag majoritet 58,8%)(v3) att de inte har *tillräckligt* hög tillit till systemets bedömning. Resultatet måste därför analyseras tillsammans med Lukyanenkos studie om tillit till AI; och samtidigt sättas i sammanhanget av den övergripande (strategiska) uppfattningen av risk och tillit som nämns i forskningsöversikten (Scharre, 2018; Claesson, 2021). Förklaringen kan då tolkas ligga i att västerländska försvarsmakter bygger sina system på att bibehålla en människa i beslutsfattandet om vapeninsatser (kontrollperspektivet). Officeren gör således denna avvägning med bakgrund av den information som finns att tillgå i ljuset av ett större strategiskt perspektiv.

En förklaring till att hellre riskera det egna förbandets liv skulle därför vara att officeren (respondenten) hade för lite information att tillgå och därför avbryter åtgärden. Detta skulle kunna beskrivas som professionsheuristik. Det vill säga beslut fattade på tidigare kunskap utan att man i detalj kan förklara varför (Lukyanenko, m.fl. 2022 sida 2005).

Studiens sista del fråga handlar om vilka mer generella attityder till användning av AI i militära operationer som uttrycks i svaren. Respondenterna uttrycker dels vikten av att det är en officer som fattar besluten genomgående i vinjetterna, särskilt i de mer riskfyllda operationerna. Officerarna visar också en viss nyfikenhet och förväntan att AI kommer att underlätta arbetet i svåra situationer. Officerarna tycks alltså anpassa och använda sin tillit och riskperception som en slags situationsbaserad heuristik för att värdera AI-systemets förslag. Detta måste tas tillvara i utvecklingen, då nyfikenhet kan skapa nya tekniska- och taktiska användningsområden. Vinjettsvaren kan dessvärre inte generaliseras till en större population men ger åtminstone stöd åt nuvarande forskning om tillit till AI (Lukyanenko, m.fl., 2022) och risk (Haerem m.fl. 2011).

Däremot finns en aspekt som jag inte riktigt räknat med: Trots att trots att en del respondenter anser sig ha relativt goda kunskaper finns också svar som visar en viss skepticism, baserad på vad jag bedömer som kunskapsbrist, mot användandet av AI. Denna generella attityd kan förklaras, med stöd av Lukyanenkos (m.fl. 2022 sida 1997-1998) studie, med att användandet av AI påverkas av personliga förkunskaper men också organisationella och kulturellt betingade erfarenheter som användandet av AI i andra sammanhang.

Detta är i sig inte nödvändigtvis negativt men kan vara en utmaning för HMT-utvecklingen som nämndes i forskningsöversikten (Warren och Hillas, 2020; Claesson, 2021) då officerarna troligen kommer att möta en allt mer digitaliserad vardag. Således behöver HMT-utvecklingen inte enbart inrikta sig på teknikanpassningar utan även utveckla användarnas systemförståelse.

Sammanfattningsvis visar resultaten i denna studie att respondenterna har en försiktig tillit till AI-system i militära operationer även i komplexa situationer där risk för mänskligt lidande och död utgör en faktor av många (om än en viktig sådan). En klar majoritet av respondenterna vill inte att AI självmant fattar beslut som innebär mänskligt lidande eller död, men som tillämpningen i vinjetterna har visat kan det finnas situationer då användandet av sådana system kan tillåtas eller motiveras, så länge det finns möjlighet att återta kontrollen över detta system (v2).

För att möjliggöra införandet av AI i militära operationer krävs dock att användarna förstår på vilka grunder som ett handlingsalternativ har genererats, dvs. black box-problematiken upplevs som hämmande. Om möjligt, bör därför XAI-aspekter beaktas i framtida AI-integreringar i Försvarsmakten.

6.2. Bidrag till professionen

Denna studie har haft som ett övergripande syfte att beskriva tilliten och riskperceptionen hos svenska officerare till att använda AI i militära operationer. Uppsatsen i sin helhet bör därför ses som ett professionsbidrag, vilket förstärks av studiens sätt att löpande beskriva utmaningar och möjligheter för framtida AI-utveckling i Försvarmakten.

6.3. Metodologiskt bidrag

Den stora vinsten med att ha genomfört studien om tillit till AI i militära operationer som en vinjettstudie och inte enbart som en traditionell enkätstudie är att respondenterna fick möjlighet att väga in ett större antal parametrar innan de gjorde valet att lita- eller inte lita på AI-systemet. Här drog jag lärdomen att trots att jag ställde kontrollfrågor så fanns det tolkningar som föll emellan vinjetten och kontrollfrågan. Hade jag gjort om studien idag hade jag försökt fånga detta med att skapa fler värderingsskalor, såsom ”håller du med om att...”

Att genomföra en vinjettstudie var intressant men gav ett stort merarbete för att analysera resultatet. Metoden kräver således en hög ambition från forskaren och förståelse för forskningsområdet för att genomföras som ett mastersarbete, exempelvis genom framtagandet av vinjetterna. Ambitionen att lyckas med vinjettstudien gjorde att Sheringham (m.fl. 2021. Sida 4.) råd och krav för en god studie, med facit i hand tolkades lite väl hårt, 1–2 vinjetter inklusive följdfrågor hade räckt. Intressant nog gav insatserna för att dölja studiens syfte resultat, vilket återspeglades i vissa svar som sökte klargöra detta eller påvisa att frågorna var fel ställda.

Vinjetterna innehöll också ett för stort antal variabler i scenariot för att alla svaren skulle kunna grupperas i ett gemensamt dataset. Således skulle kritik mot en vinjettstudie vara att det är svårt att operationalisera studiens alla variabler, detta krävde stor tankemöda under framtagandet och metodens tillämpning i framtiden bör beakta just vinjettstudiens operationalisering. Här rekommenderas den som vill göra en vinjettstudie att överväga två olika vinjettyper istället, såsom en texttolkande (likt denna) och en bildtolkande vinjett.

Tolkningen av resultatet innebar mer än att studera statistik med hjälp av program som SPSS, vilket också krävde en avvägning av vilket resultat som presenterades i studien. För mycket statistik skulle missa de viktiga underliggande nyanser som återfanns i fritextsvaren. Detta skulle säkerligen kunna bli lättare i framtiden om det anses etiskt försvarbart att använda AI-baserade hjälpmedel för att strukturera och analysera svaren, som Nvivo. Jag kände dock att det var en för stor risktagning att använda dessa system utifrån forskningsetiska frågor såsom vart material lagras, vem har tillgång till studiens rådata etc.

Sammanfattningsvis kan sägas att vinjettstudiernas stora fördel för krigsvetenskapen är att frågorna går att projicera mycket närmare verklighetens komplexitet som den ser ut idag. snarare än att enbart forska på detaljer av det som varit. Nyanserna i svaren möjliggör förståelse för den mångbottnade problembild som återfinns i krigets natur. Nackdelen är att de riskerar att bli för stora empiriskt. Fler krigsvetenskapliga studier skulle tjäna på att utformas som vinjettstudier eller som vinjettexperiment. Studien bjuder också in till möjligheten att skapa fler multidisciplinära studier, som att forska på både ledarskap, etik och filosofiska aspekter av krigföringen.

6.4. Fortsatt forskning

Artificiell intelligens i militära operationer har stor potential att studeras vidare. Bland annat rekommenderas fortsatt forskning utgå ifrån denna studie för att på en högre nivå (ekologisk studie) jämföra olika nationers officerskårer, samt att identifiera underliggande orsaker till eventuella skillnader i kontrollbegreppet som jag valde att utelämna. Vidare föreslås mer forskning inrikta sig på användarnas etiska och moraliska grunder kopplat till användandet av AI på andra nivåer, såsom taktiska aspekter eller vad som händer om AI används i framtagande av militärstrategiska handlingsalternativ eller optioner. I Appendix 1 finns flera uppslag till fortsatta studier på olika nivåer, bland annat om förståelse och officersutbildningarnas omhändertagande av AI och disruptiv teknik.

6.5. Avslutande reflektion

En viktig detalj som inte fick utrymme men som jag tycker bör belysas är att militariseringen av AI är påbörjad och en oundviklig framtid att förhålla sig till. *Oavsett* om forskningen i väst varnar för det kommer försvarsmakter världen över agera utifrån direktiv från sina regeringar. Således, om inga begränsningar mot att utveckla militära AI-system sker på en global nivå står vi inför samma dilemma som när kärnvapnen först utvecklades.

Därför slutar denna uppsats där den började, i Tegmarks öppna brev med en förmaning om att vara försiktig med en kapprustning inom AI innan vi kan kontrollera dess effekter. För vad händer exempelvis om två stridande parter använder samma AI? Ligger då livets öde hos maskinen eller människan?

7. Referensförteckning:

- Antebi, L. (2021) "Artificial Intelligence and National Security in Israel", *INSS*, s. 2021–2023.
- Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*. 1:a uppl. Oxford: Oxford University Press, Incorporated.
- Brehmer, B. (2013) *Insatsledning : [ledningsvetenskap hjälper dig att peka åt rätt håll]*. 1. uppl. Stockholm: Försvarshögskolan.
- Claesson, M. (2021) *Vägval : framtiden för svensk säkerhet*. Redigerad av Z. Carlander. Stockholm: Ekerlids förlag.
- Derian, J. Der och Rollo, S. (2023) "Quantum warfare", *Routledge Handbook of the Future of Warfare*, s. 319–329.
- Esaiasson, P. (2017) *Metodpraktikan : konsten att studera samhälle, individ och marknad*. Femte upplagan. Redigerad av M. Gilljam m.fl. Stockholm: Wolters Kluwer.
- Fels, E. (2023) "Prospects of great power rivalry: Escaping the tragedy?", *Routledge Handbook of the Future of Warfare*, s. 377–388.
- Finlan, A. (2021) "The shape of warfare to come: a Swedish perspective 2020–2045", *Defense and Security Analysis*, 37(4), s. 472–491.
- Forden, G. (2000) "False alarm, nuclear danger", *IEEE Spectrum*, 37(3), s. 31–39.
- González, R.J. (2023) "Militarising Big Tech The rise of Silicon Valley's digital defence industry STATE OF POWER 2023 STATE OF POWER 2023", s. 1–14.
- Hadlington, L. m.fl. (2023) "The use of artificial intelligence in a military context: development of the attitudes toward AI in defense (AAID) scale", *Frontiers in psychology*, 14, s. 1164810–1164810.
- Haerem, T. m.fl. (2011) "Do military decision makers behave as predicted by prospect theory?", *Journal of Behavioral Decision Making*, 24(5), s. 482–497.
- Holland Michel, A. (2020) *The Black Box, Unlocked: Predictability and Understandability in Military AI*. Geneva, Switzerland.
- Husain, A. (2021) "AI is Shaping the Future of War", *Source: PRISM*, 9(3), s. 50–61.
- Hynek, N. och Solovyeva, A. (2022) *Militarizing Artificial Intelligence, Militarizing Artificial Intelligence*.
- Innes, J.M. och Morrison, B.W. (2021) "Artificial intelligence and psychology", *The Routledge Social Science Handbook of AI*, (April), s. 30–57.
- Johnson, J. (2023) "Automating the OODA loop in the age of intelligent machines: reaffirming

the role of humans in command-and-control decision-making in the digital age”, *Defence Studies*, 23(1), s. 43–67.

Kahneman, D., Tversky, A. och Tversky ’, A. (1979) ”Prospect Theory: An Analysis of Decision under Risk E C O N OMETRICA I C I VOLUME 47 MARCH, 1979 NUMBER 2 PROSPECT THEORY: AN ANALYSIS OF DECISION UNDER RISK”, *Source: Econometrica*, 47(2), s. 263–292.

Kania, E.B. (2019) ”Chinese Military Innovation in the AI Revolution”, *The RUSI journal*, 164(5–6), s. 26–34.

Kissinger, H. (2021) *The age of AI: and our human future*. Redigerad av E. Schmidt och D. Huttenlocher. London: John Murray.

Lidskog, R. (2020) ”Samhället utmanat? Artificiell intelligens och sociologisk kunskap”, *Sociologisk Forskning*, Årgång 57.(2), s. 99–120.

Lukyanenko, R., Maass, W. och Storey, V.C. (2022) ”Trust in artificial intelligence: From a Foundational Trust Framework to emerging research opportunities”, *Electronic Markets*, 32(4), s. 1993–2020.

Maas, M.M. (2019) ”How viable is international arms control for military artificial intelligence? Three lessons from nuclear weapons”, *Contemporary Security Policy*, 40(3), s. 285–311.

Nyholm, S. och Smids, J. (2016) ”The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem?”, *Ethical Theory and Moral Practice*, 19(5), s. 1275–1289.

Png, M.T. (2022) *At the Tensions of South and North: Critical Roles of Global South Stakeholders in AI Governance, ACM International Conference Proceeding Series*.

Raska, M. (2021) ”The sixth RMA wave: Disruption in Military Affairs?”, *Journal of Strategic Studies*, 44(4), s. 456–479.

Roberts, H. *m.fl.* (2021) ”The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation”, *AI & society*, 36(1), s. 59–77.

Rosert, E. och Sauer, F. (2019) ”Prohibiting Autonomous Weapons: Put Human Dignity First”, *Global policy*, 10(3), s. 370–375.

Scharre, P. (2018) *Army of none : autonomous weapons and the future of war*. First edition. New York: W. W. Norton & Company.

Sheringham, J., Kuhn, I. och Burt, J. (2021) ”The use of experimental vignette studies to identify drivers of variations in the delivery of health care: a scoping review”, *BMC Medical Research Methodology*, 21(1), s. 1–17.

- Stefan Axberg, Kent Andersson, Martin Bang, Nils Bruzelius, Peter Bull och Per Eliasson, Marika Ericson, Mikael Hagenbo, Gunnar Hult, Eva Jensen, Hans Liwång, Lars Löfgren, Martin Norsell, Åke Sivertun, C-G Svantesson, B.V. (2013) *Lärobok i Militärteknik, vol 9: Teori och metod*. Försvarshögskolan.
- Sullivan, Y., de Bourmont, M. och Dunaway, M. (2022) "Appraisals of harms and injustice trigger an eerie feeling that decreases trust in artificial intelligence systems", *Annals of operations research*, 308(1–2), s. 525–548.
- Taylor, R.D. (2020) "Quantum Artificial Intelligence: A "precautionary" U.S. approach?", *Telecommunications policy*, 44(6), s. 101909.
- Tegmark, M. (2017) *Life 3.0: being human in the age of artificial intelligence*. New York: Alfred A. Knopf.
- Tversky, A. och Kahneman, D. (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty Author (s): AMOS TVERSKY and DANIEL KAHNEMAN Published by : Springer Stable, *Springer*, 5(4), s. 297–323.
- Wakker, P.P. (2010) *Prospect theory: for risk and ambiguity*. Cambridge: Cambridge University Press.
- Warren, A. och Hillas, A. (2020) "Friend or frenemy? The role of trust in human-machine teaming and lethal autonomous weapons systems", *Small Wars and Insurgencies*, 31(4), s. 822–850.
- Wasilow, S. och Thorpe, J.B. (2019) "Artificial Intelligence, Robotics, Ethics, and the Military: A Canadian Perspective", *The AI magazine*, 40(1), s. 37–48.

APPENDIX 1

I detta appendix redovisas enkäten i sin grundform, utan vissa interaktiva delar så som svarsboxar och rullister av formateringskäl. Respondenternas svar är inte bifogade av forskningsetiska skäl.

Välkommen att delta i en studie av beslutsfattning i framtida operationer

Hej och tack för att du deltar i min vinjettstudie om beslutsfattning i framtida operationer! Enkäten bygger på frivillighet och du har möjlighet att närsomhelst avbryta studien. Den går lika bra att svara i datorn som i surfplatta eller telefon men vinjetterna har en del text att läsa så jag rekommenderar inte att göra den i din telefon. Vinjetterna är till största del fiktiva men en av vinjetterna är baserad på en verklig händelse. Frågorna resonerar en del kring användandet av AI. Här avses sådan AI som bygger vidare på den typen av AI-tekniker/modeller som finns idag (tänk chatGPT, Bard, Copilot).

Det är enbart jag som skapat enkäten som har tillgång till dina svar och de kommer enbart presenteras i mitt masterarbete och presentationsunderlag kopplat till denna. Dina svar kan inte kopplas ihop med dig som individ men vissa fritext-svar kan komma att citeras i arbetet. Frågorna kan te sig enkla i vissa avseenden, vänligen svara på dem ändå utan att tolka in något annat syfte då svaren kommer att jämföras mot ett annat material och har anpassats för att vara jämförbara. Svaren bygger också på att du svarar utifrån dina egna bedömningar, inte vad du tror att någon annan förväntar sig. Enkäten är interaktivt uppbyggd och frågor kan dyka upp eller försvinna beroende på ditt svar. Enkäten tar ca 15 minuter att besvara.

DEL 1 Allmänna frågor Godkänner du att vara med i studien på det sätt som beskrivs ovan?

- Ja
 Nej

Vilket åldersspann tillhör du?

- 25-35
 35-45
 45-55
 55-65

Vilket truppslag eller stridskraft har du mest erfarenhet ifrån?

- Armén
 Flygvapnet
 Marinen
 Led/und
 Rymd/cyber

På en skala 1-5 hur värderar du din teknikmognad i jämförelse med dina officerskollegor?

- 1 = mycket lägre förmåga
 2 = något lägre förmåga
 3 = samma förmåga
 4 = något högre förmåga
 5 = mycket högre förmåga

Hur ofta använder du AI-stödda hjälpmedel i vardagen, exempelvis Chat-GPT eller röststyrda högtalare?

- Aldrig
 någon gång i månaden
 1-2 gånger i veckan
 3-4 gånger i veckan
 Varje dag

Har du använt AI i din ordinarie tjänst? (ej i skolmiljö)

- Ja
 Nej

DEL 2: VINJETTER

Förutsättningar:

Du befinner dig ca 10–15 år längre fram i tiden och du är chef över en stridsgrupp med både armé-, flyg- och marinstridskrafter till förfogande. Sedan en tid tillbaka är delar av förbandet på plats i en av de baltiska staterna medan delar av förbandet är baserat i Sverige (vissa marin och flygstridskrafter). Sverige är medlem i NATO men organisationens huvudfokus ligger kring den polska gränsen mot Ukraina och Belarus där situationen blir alltmer spänd med ovanliga truppansamlingar och flera allvarliga incidenter. Det primära uppdraget för ert styrkebidrag är därför att öka säkerheten och närvaron i de baltiska staterna under artikel 3, med relativt långtgående mandat att agera offensivt inom ditt AOR.

Omfattande cyber-angrepp har genomförts mot vital it-infrastruktur vilket påverkar sambandet med Sverige och NATO-förband negativt och ni har begränsad tillgång till era ledningsstödssystem. Det enda system som fungerar felfritt är Försvarsmaktens nya AI-beslutsstödssystem som utvecklats under 15 års tid, Det verkar ha haft en starkare kryptering och därmed inte kunnat forceras.

Vinjett 1

Sedan en dryg vecka tillbaka ligger ett av era större fartyg förtöjd på grund av motorhaveri i en större hamnstad. I staden har stora desinformationskampanjer om er närvaro florerat en tid, tonen i de sociala medierna har sedan i helgen tagit en ny hårdare vändning och anklagelser mot en grupp sjömän med namn och bild för att ligga bakom ett övergrepp mot minoritetsbefolkningen i staden, vilket har lett till att en demonstration har tagit fart. Lokalpolisen hanterar situationen på platsen och ni har endast en förstärkt bevakning i hamnen.

Samtidigt får ni rapporter om att ett stridspar som bevakat fartyget i hamnen har övermannats och förts bort. De är relativt oskadda och visas upp på en film i sociala medier men samtidigt ställs ett krav på att utlämna de soldater som pekats ut för övergreppen i en utväxling.

Senare på kvällen redovisar din stab inför dig och dina rådgivare vad de vet om rådande läge och föreslår två handlingsalternativ för att lösa situationen. Ett offensivt med kinetiskt våld samt ett offensivt men med icke-kinetiskt våld.

Det kinetiska alternativet bygger på en fritagningsoperation med specialförband i ett förortsområde där gisslan förväntas hållas. Det kommer eventuellt att lösa gisslansituationen, din stabschef tycket att detta vore det bästa alternativet men ert AI-system avråder från detta då tidigare insatser som liknat denna visat att denna åtgärd skulle vinklas negativt och anses som provokativ mot minoritetsbefolkningen med konsekvenser som ej kan beräknas. Risken för att någon part i en sådan operation dödas beräknas till 50–75%.

Det icke-kinetiska alternativet innebär en offensiv kombinerad cyber- och informationsoperation för att dels förstöra de element (servrar och datorer) som används i desinformationskampanjen och parallellt genom olika former av digitala intrång lokalisera och följa gisslantagarnas aktiviteter, risken är dock att gisslan kan hinna flyttas. Ert AI-system anser att det icke-kinetiska alternativet är det mest lämpade alternativet då det har samma sannolikhet att lyckas men mindre risk för att någon part dödas (under 10%).

Hur värderar du AI-råden och vilket beslut tar du?

- Jag litar inte på AI-systemet och beslutar mig för det offensiva kinetiska alternativet.
- Jag litar på AI-systemet och beslutar mig för att genomföra det offensiva icke-kinetiska alternativet.
- Om annat, specificera

Hade du beslutat annorlunda om AI-systemet förordade det andra alternativet?

- Nej, jag litar inte på att AI-system kan göra sådan rekommendationer.
- Nej, jag vet inte hur den kommer fram till sina rekommendationer.
- Ja
- Om annat, specificera

Vinjett 2

Händelseutveckling:

En tid efter att situationen i Hamnstaden lugnat ner sig och era soldater kunnat lokaliseras och fritas har situationen i en annan del närmare gränsen urartat. Påstådda separatister har sedan en månad stridit mot lokala arméförband. Separatisterna har ockuperat en mindre stad och deklarerat att man avser marschera mot huvudstaden, ta över makten och slänga ut NATO ur landet. Ditt förband har beordrats att omhänderta situationen från att urarta och skära av separatisternas möjligheter att avancera. Samtidigt är ni inne i en rotationsperiod och arméförbanden samt en stor del av materielen som varit i landet ska lösas av inom kort. De marina styrkorna förbereder och skyddar överskeppning av soldater och materiel från en svensk hamn inom några dygn.

Er stab har genom AI-beslutstödsystemet har genererat tre möjliga handlingsalternativ för att hantera utvecklingen:

Det första alternativet innebär att invänta utvecklingen och låta den lokala armén fortsätta binda separatisterna medan förbandet genomför rotationen enligt planen, för att sedan genomföra operationen så snart förbandet är på plats. Din stabschef anser att detta är det bästa alternativet, både ur en personalpolitisk och ekonomisk synvinkel. AI-systemet räknar de initiala riskerna för egna förband till 0% men den lokala armén förväntas förbruka 50–75% av sina tillgängliga resurser i detta alternativ.

Det andra alternativet föreslår att behålla både det avgående och pågående förbandet i upp till två månader för att förbereda ett försvar med goda förbandsreserver. Det alternativet kan på sikt innebära att båda förbanden binds i strid om separatisterna bryter sig ut, och då finns inga reserver redo hemma i Sverige. Det ger dock en starkt avskräckande effekt.

Det tredje alternativet bygger på att verka proaktivt, och innebär att det ankommande förbandet får i uppgift att skyndsamt påbörja försvar av huvudstaden medan det förbandet som redan är på plats framgrupperar och understödjer den lokala arméns återtag av den ockuperade staden. Insatsen har goda chanser att lyckas då ni tillsammans med den lokala armén skulle ha ett mycket stort övertag i personal och kvalificerade resurser.

AI-beslutstödssystemet förordar det tredje alternativet då det har stora fördelar. Det ger en positiv bild av förbandet hos lokalbefolkningen. Samtidigt visar Sverige och NATO handlingskraft och verkar avskräckande. Vilket kan lugna ner situationen i regionen. Det är dock inte riskfritt och sannolikt kommer detta innebära materiella och personella förluster för er, motståndaren och för de civila som bor kvar i byn.

Hur värderar du situationen och vilket beslut förordar du?

- Jag väljer att rotera hem förbanden, jag vill inte riskera personal eller ekonomi för en situation som jag inte kan förutse.
- Jag väljer att behålla båda förbanden, jag tror mer på ett defensivt agerande utan att riskera egna förband.
- Jag genomför den offensiva operationen, Jag litar på AI-systemets beräkningar och motiv, jag är beredd att ta riskerna då det ger stora fördelar att agera proaktivt.
- Om annat, specificera

Hade du beslutat annorlunda om AI-systemet förordade något av de andra alternativen?

- Nej, jag litar inte på att AI-system kan göra sådan rekommendationer och tog beslutet av andra anledningar
- Nej, jag vet inte hur den kommer fram till sina rekommendationer.
- Ja, för att jag litar på att AI i framtiden kan förutse händelser och risker som vi missat.
- Om annat, specificera

Låt oss gå vidare med planen att genomföra den offensiva operationen. Efter en veckas strider visar det sig att fienden forskansat sig hårdare än tänkt. Både ni och den lokala armén har fått personella förluster och AI-beslutsstödsystemet föreslår att sätta in autonoma vapensystem som kan söka upp och oskadliggöra fienden utan att riskera fler egna förluster. Hur värderar du en sådan situation?

- Jag ser det som ett självklart alternativ, riskerna för egna förluster motiverar det.
- Jag nekar användandet av sådana system, en människa måste ta beslut om vapeninsats
- Om annat, specificera

Vinjett 3:

Tänk dig att du är fartygschef och patrullerar vattnen utanför ett terrordrabbat område. En stark nation i området stödjer de terrorister som utfört attacker mot civila och militära fartyg. Terroristerna har således tillgång till ett stort antal kvalificerade resurser.

Ett varningsmeddelande från er radar visar plötsligt att ett oidentifierat flygplan närmar sig er position i hög hastighet. Fartygets nya AI-system ombord är inställt på att avfira om hotet bedöms tillräckligt högt. Ni hör att en luftvärnsrobot avfyras av systemet och du har max 2 minuter på dig att avbryta när roboten väl är i luften. Du vet samtidigt att luftrummet inte är stängt för civila flygplan och tittar därför på AI systemets statistik. Den anger att det är med 65% sannolikhet ett fientligt flyg då det kommer från den terrorstödjande staten.

Om du låter robotarna gå mot målet så hinner de verka mot det innan farkosten kan påverka er men du kan inte identifiera vad det är för något.

Avbryter du robotarna finns det bara skyddet från er allmålskanon kvar vilket utgör en stor risk mot fartyget och besättningen men ni kommer hinna identifiera flyget.

Vad gör du?

- Jag litar på AI-systemet och låter roboten fortsätta, det är värt riskerna.
- Jag väljer att avbryta och identifiera flygplanet, det är värt riskerna.

DEL 3 AI i Försvarsmakten

Anser du att vi är inne i ett tekniksprång där AI kommer vara en del av den militära professionen?

- Ja
- Nej

Tycker du att:

	NEJ	JÄ
AI ska kunna användas för att optimera militära logistikflöden:	<input type="checkbox"/>	<input type="checkbox"/>
AI ska tillåtas som hjälpmedel för att planera militära operationer?	<input type="checkbox"/>	<input type="checkbox"/>
AI ska kunna ge order direkt till enheter som är i fara?	<input type="checkbox"/>	<input type="checkbox"/>
AI ska själv kunna fatta beslut om offensiva åtgärder men som inte leder till mänskligt lidande eller död? (ex genomföra informations- eller cyberoperationer)	<input type="checkbox"/>	<input type="checkbox"/>
AI ska själv kunna fatta beslut som innebär lidande eller död?	<input type="checkbox"/>	<input type="checkbox"/>

Utifrån det läge ni nu planerar inom LGO-kursen, tror du resultatet hade blivit bättre eller sämre om Försvarsmakten hade ett eget AI-verktyg som kunde generera alternativ utifrån Försvarsmaktens förutsättningar likt ChatGPT, Bard eller Copilot kan?

- Ja
- Nej
- Vet ej

Tror du att våra motståndare kommer att utveckla helt autonoma vapensystem som själva tar beslut om verkanseld (exempelvis sjö-, land- eller luftburna drönare)?

- Ja
- Nej
- Vet ej

Tror du att AI kan utveckla din stridskrafts- eller försvarsgrens förmågor?

- Ja
- Nej
- Vet ej

Anser du att Försvarsmakten måste ändra sina beslutsprocesser för att implementera AI-stöd på ett bättre sätt?

- Ja
- Nej
- Vet ej

Tror du att du måste ändra ditt arbetssätt på grund av ny teknik, så som AI?

- Nej
- Ja
- Vet ej

Anser du att dagens officersutbildningar (OP/HOP) omhändertar framtida påverkan av ny teknik så som AI?

- Inte alls
- Till del men det är för abstrakt och inte tillämpbart
- Till del men kan utvecklas
- Fullt tillräckligt och på rätt nivå för att vi ska förstå utmaningarna
- Mer än nödvändigt för de flesta, vi borde lägga mindre tid på sådan utbildning.

Avslutning Skulle du vilja tillägga något annat? Var vinjetterna och frågorna rimliga eller missades något?

- Nej tack
- Ja jag skulle vilja tillägga: