Automation as an Intelligent Teammate: Social Psychological Implications

José Kerstholt, Jonathan Barnhoorn, Tom Hueting & Lotte Schuilenborg
TNO
PO Box 23
3769 ZG Soesterberg
THE NETHERLANDS

Jose.Kerstholt@tno.nl

ABSTRACT

In order to increase the effectiveness of NATO operations with flexible and robust human-autonomy teams, insight is needed into the psychological mechanisms underlying teaming with artificial agents. We conducted a literature scan to provide an overview of current knowledge on factors affecting both objective and subjective outcomes of human-autonomy teaming. The main results from our study are that current research mainly addressed factors affecting overall task performance, like situation awareness, and neglected relational aspects and human appraisal. However, social outcomes ensuing from two-way interactions will become more and more important with increased agency of AI. We therefore argue that in order to fully utilize the potential of autonomous systems, more insight is needed into (unconsciously operating) relational aspects, and how they interact with task-related factors.

1.0 INTRODUCTION

Lieutenant Bob is performing an overt reconnaissance mission in a hilly rural area together with his soldiers. The squad uses multiple motorised vehicles to transport itself. At some point in time a swarm of unmanned aerial vehicles (UAVs) that accompanies them locates enemy personnel. The swarm, communicating through an artificial agent called Smart Bud, transmits the enemy positions to all friendly personnel and systems in the area. This includes a 60mm mortar crew which immediately starts preparations to potentially engage the enemy. During this procedure, Smart Bud informs Lieutenant Bob of the 60mm mortar crew's selected targets and advises Lieutenant Bob to cancel this plan on the basis of a disproportional amount of expected collateral damage. Instead, re-planning is advised. Based on his prior combat experience and the challenging operational situation at hand, Lieutenant Bob disagrees with Smart Bud's advice. He informs Smart Bud about his concerns, resulting in a short dialogue about pros and cons of both courses of action.

Artificial Intelligence (AI) has brought about major changes to almost every aspect of our lives and although expectations differ wildly, it may reach a point where it equals or even surpasses human intelligence (Grace, Salvatier, Dafoe, Zhang, & Evans, 2017). Approaching this point requires rethinking our relationship with technology, as AI is going beyond intelligent (decision) support by augmenting human knowledge and skills and facilitating unprecedented capabilities. The scenario above illustrates a situation in which AI (the Smart Bud) is on equal footing with the human when making consequential choices. In other words, Smart Bud has become a team member with a high level of agency (Nyholm, 2017). What would the dialogue between the Smart Bud and Bob look like regarding arguments being exchanged, or even more advanced, regarding social-skills being applied to convince the other team member of the supremacy of one's position? Would it affect the actual decision being taken?

Even though military decision-making is preferably considered as a rational process (e.g., Boyd, 1995), in reality social-psychological mechanisms like team cohesion are central to team functioning (Salas, Sims, & Shawn Burke, 2005). Thus, it may well be that an interaction between a human and AI as described above triggers such psychological mechanisms, which therefore need to be taken into account when designing AI. How will humans react to team members being replaced by AI? Will it affect their performance? Their motivation? Will we see AI as an actual team member or will we always see it as a 'support' system? The present paper tries to explore some issues related to this new area of AI.

2.0 STATE OF THE ART

The present paper reports the results of a literature review which was meant to identify knowledge gaps with regard to (social-) psychological factors that are relevant to teaming with AI that has high levels of agency and assuming the role of a team member. This is often referred to as a human-autonomy team (HAT). We have organised our findings in terms of an IMOI framework (input-mediator-outcome-input) postulated by Ilgen, Hollenbeck, Johnson, & Jundt (2005), which is commonly used in the literature on team effectiveness (see Figure 1). The model looks at the processes of a team as an integrated system with the organisation, the team itself and the individual members serving as input. The processes that happen during the teamwork serve as mediators, and multiple criteria count as outcomes of the team work (i.e. team satisfaction or taskwork outcomes). In addition, there are feedback loops throughout the process (new input) so that outcomes, for instance, can have an effect on the mediators and input. For instance, team satisfaction as an output can, in a subsequent feedback loop, have influence on inputs or mediators (e.g. team cohesion). Despite being used extensively in the human-human teaming literature (Ilgen, Hollenbeck, Johnson, & Jundt, 2005), this model has also been used for a similar but somewhat less exhaustive analysis by You and Robert (2017) as well as by Stowers et al. (2017) in the HAT literature. This paper adds to this discussion by a stronger focus on the social-psychological mechanisms in HATs which have been largely ignored in previous work. In addition, knowledge gaps related to these mechanisms are identified for the purpose of future research.

This paper is structured along the lines of the IMOI framework. First, we describe the outcomes relevant to HAT. Second, we describe the mediators and how they affect the outcomes, especially as compared to human-human teaming, and lastly, we describe the inputs. The model is described in reverse order to clearly show the feedback loops and the gaps in the literature.

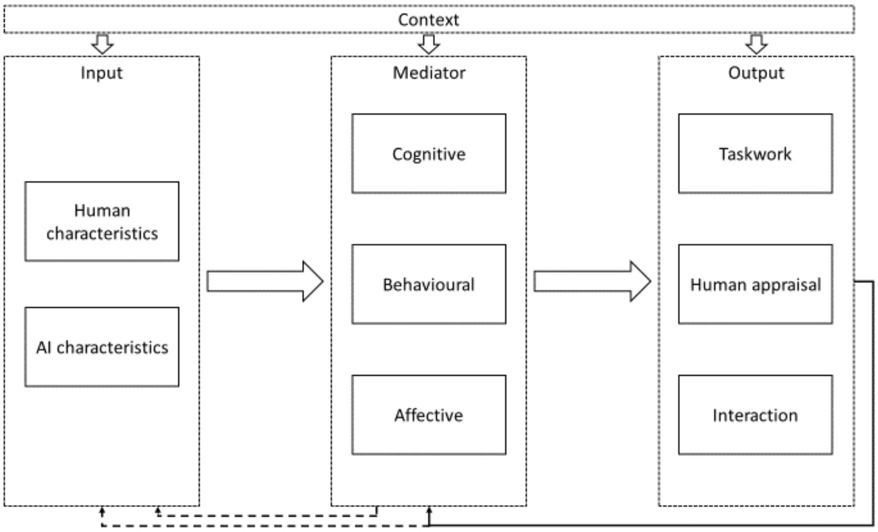

Figure 1: Input-Mediator-Output (IMO) team Effectiveness Framework

2.1 Outcomes

The model distinguishes three classes of outcomes: taskwork outcomes (relatively objective variables indicating how well the task is completed, e.g., errors made), human appraisal, (e.g., a person's satisfaction), and interaction outcomes (that inherently emanate from working in a team, e.g., calibrated trust).

2.1.1 Taskwork outcomes

A general finding from our literature review is that most studies have focused on task performance, or the completion of goals and sub goals. Task performance is often measured in terms of time needed to perform a task, error rate, quality, or efficiency (e.g., number of mouse clicks needed to reach a certain level of performance (Clare, Cummings, & Repenning, 2015)). One study into adaptive aiding in military HATs, for example, measured mission completion time, detection performance and classification performance (Visser & Parasuraman, 2011). Another study using a multi-UxV planning system 'IMPACT' measured the rate at which operators correctly accepted or rejected plans proposed by an intelligent assistant (Mercado et al., 2016). Other measures used include target detection performance in a gunnery task with an intelligent aided target recognition system (Chen & Terrence, 2009), and area coverage in a surveillance task collaborating with a swarm of UAVs (Clare et al., 2015).

Another observation, related to task work outcomes, is that studies on AI in a military context are primarily anchored in the current operating paradigm of fighting and defeating enemies (Spiegeleire, Maas, & Sweijs, 2017). As such, AI is seen as an extension of technological development to increase the effectiveness of operations, with lower risks and higher speed. Think, for example, of algorithms to interpret imagery from drone surveillance feeds, remotely-piloted systems, and AI-coordinated swarms. This focus may explain the relative abundance of studies on optimising task performance.

2.1.2 Human appraisal

Human appraisal is deeply engrained into the human-human teaming literature (Ilgen et al., 2005). Examples of measures of human appraisal include various forms of satisfaction (e.g. with the team, with the outcome or with the collaboration), attribution and agency, coping/mental resilience, acceptance and self-efficacy. In general such research is relatively sparse in the context of HAT, even though some of these measures have been studied.

Some studies have, for example, focused on topics like technology acceptance, the attribution of blame and self-efficacy. It can, for instance, be measured how (much) blame is attributed to humans and AI for an action with a negative (unintended) effect (Malle, Scheutz, Arnold, Voiklis, & Cusimano, 2015a). Closely related are studies on the attribution of moral permissibility which can be measured by the degree to which people find the actions of the AI or the human justifiable and/or permissible (Malle et al., 2015a). Self-efficacy has been studied in relation to trust, showing that a person's confidence in his or her own skills predicts the extent to which he relies on the automation (e.g., Prinzel, 2002).

2.1.3 Interaction outcomes

A number of outcomes are related to the interaction within the team. For instance, the effectiveness of communication and, related to this, situation awareness (SA) on the individual, team, or even organisational level (Stanton, Salmon, Walker, Salas, & Hancock, 2017). Another outcome of the team interaction processes is the shared mental model of, for instance, the task, goals and member capabilities the team has developed during task execution. An optimally shared mental model results, according to these studies, in a shared and correct view of how to reach objectives (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000). A correct (shared) mental model also results in appropriately calibrated trust in the AI, which is another interaction outcome that has received extensive attention (Schaefer, Chen, Szalma, & Hancock, 2016). Overall, interaction outcomes seem to have received a fair share of attention in HAT research.

2.1.4 Gaps with regard to outcomes

With regard to outcomes that have been addressed in current research, we concluded that least attention has been paid to human appraisal, that is, subjective evaluations of team performance and functioning. There has been a focus on overall task performance in current HAT studies, as the main aim seems to be to improve overall team outcome by optimally combining the qualities of human and non-human agents (e.g., Onnasch, Wickens, Li, & Manzey, 2014). In such a context important questions are, for example, how to divide tasks across actors by taking workload into account (adaptive automation), or which knowledge AI needs in order to optimally support a human agent. A broader view, however, would be to consider factors like motivation of human team members and human mental resilience. This broader view would then assume the AI to be more than a decision-support system, but a real team member.

Going back to our initial scenario for example, it would be necessary to know how lieutenant Bob evaluates the arguments of Smart Bud, whether he trusts and follows them, and how the team would continue their interaction on both task and social level after experiencing negative consequences due to wrong decisions (resilience).

2.2 Mediators

In order to describe the mediators, we made a distinction between cognitive and behavioural factors on the one hand, and affective factors on the other.

2.2.1 Cognitive and behavioural

Many cognitive processes and behaviours have been investigated in terms of their effect on human-autonomy teaming. These range from the mental models that individuals develop of systems, to cognitive workload, to the extent to which operators tend to monitor system performance.

Mental models are frameworks that individuals construct in order to generate predictions and support their understanding of a system's or a person's drivers, abilities and behaviour. Mental models are constructed and continuously updated by translating observations to internal representations. They are cognitive short-hands that provide guidance for interactions, without necessarily being accurate reflections of the outside-world (Cooper, Reimann, & Cronin, 2011). In HATs, the mental model that a human operator has of an AI determines the capabilities and reliability he expects from it. This way, mental models can affect many other mediators and almost all possible outcomes of a HAT. For instance, according to a recent meta-analysis, mental models are an important antecedent of trust (Schaefer et al., 2016).

Specifically, the mental model that the human has of the AI (and vice versa), regarding the expected level of knowledge, is of great influence on the style and effectiveness of communication. Previous research has identified that in order for communication to be effective there has to be a certain level of mutual understanding, or common ground, between sender and receiver (Lee et al., 2005). It has been found that people formulate their messages according to their estimate of the knowledge shared between them and the receiver. This is done intuitively between humans, albeit sometimes incorrect and/or based on preconceptions. For effective communication to occur between humans and AI, the AI has to support this as well. This means that the human must be able to gauge the level of knowledge the AI has, but the AI must also be aware of the knowledge the human has, as to avoid talking past each other.

2.2.2 Affective

Team members' motivation expands beyond attaining overall task goals, as the main drivers of military motivation are affective in nature and concern the relation with other team members (Wong, Kolditz, Millen & Potter, 2003). Team identification is important for a team to function efficiently, meaning that all members of the team see themselves as part of the team. Team identification can be created through interdependence, where the accomplishment of one individual relies on that of a teammate. Creating team interdependence can have an effect on performance and on human appraisal outcomes. For instance, Walliser, Mead, and Shaw (2017) found that human agent teams that were structured as a team through interdependence had higher performance, and higher affect then when they were not structured as a team, that is, team identity was not created. In the field of autonomous synthetic characters (i.e. autonomous players in for instance Role Playing Games) some research has also been done into team identification. Prada and Paiva (2009) showed that when social skills where programmed such that the autonomous characters exhibited behaviours consistent with the group's composition, context and structure, team identification was higher.

Trust is one of the affective mediating mechanism that has received most attention in research on working with automated or autonomous systems. The importance of this mechanism originates from the finding that trust is a strong predictor of how systems are used. For instance, too much trust results in misuse through overestimation of system capabilities, potentially leading to dangerous situations (Baker & Keebler, 2017). Undertrust, on the other hand, results in disuse of systems, potentially leading to increased workload for the human, and reduced efficiency. An extensive literature has investigated the antecedents of trust, summarized in a meta-analysis by Schaefer, Chen, Szalma and Hancock (2016). Predictors of trust include a person's age, his or her understanding of the automation, the appearance and anthropomorphism of the system, and many others. Models are under development to gain insight into how trust changes over time (e.g., Gao, Clare, Macbeth, & Cummings, 2013). Furthermore, researchers have experimented with priming participants who tend to overtrust systems (in this case, gamers), which resulted in successful calibration of trust and improved performance of the human-machine team (Clare, Cummings, & Repenning, 2015). Yet other researcher propose general models for when and how to repair trust in human machine interaction (de Visser, Pak, & Shaw, 2018).

Another important affective mediator is our perception of how our team members, in this case the AI, sees us. Even though only few studies investigated this factor, it may have profound effect on team functioning. For instance it was found that rejection by the robot after playing a game ("That was boring! I don't want to play with you again") significantly lowered self-esteem of the participant. Robot appraisal ("That was fun, I would like to play with you again some time.") on the other hand, had no influence on self-esteem (Nash et al., 2018).

2.2.4 Gaps related to mediators

A number of topics have received relatively little attention. First, when considering the functioning of HATs in a military context, we identified a gap regarding research on the human's mental model of the AI's ethical awareness, which would contribute to a correct perception of meaningful human control (Ekelhof, 2015). Meaningful human control entails the notion that humans rather than computers should remain in control of, and be morally responsible for, relevant decisions about (lethal) military operations. Second, the effectiveness of HATs is not only induced by analytic processes such as optimizing (shared) mental models but also by less rational processes such as when robot rejection lowers self-esteem. With the exception of trust, we see that such affective and irrational mediators have often been overlooked in the HAT literature. Research focusses on how we as humans view AI, but it does not yet envision AI as a team member triggering affective processes within a team. As team identification is an important prerequisite for accurate team functioning more research is needed on how to design socially skilled AI without jeopardizing task performance.

In our scenario above in which lieutenant Bob is reasoning with his Smart Bud about the best course of action, it would, for example, be interesting to know how Bob would actually decide as a function of the type of dialogue: what if Smart Bud uses some kind of persuasion, or affective non-verbal signs like a frown, or natural speech rather than text? All these factors will have large effects on the information processing of Bob and consequently on the accuracy of the decisions being made.

2.3 Input

2.3.1 Human characteristics

Literature on working with automated and autonomous systems provides many insights into individual human differences that are known to affect the way in which HATs function. Without being exhaustive, we describe the most relevant variables.

Attentional control is one's ability to focus and shift attention in a flexible manner. The literature on the effect of individual differences in attentional control on the functioning of HATs is summarised by Chen & Barnes (2014b). Attentional control was identified as one of the most important abilities that affect performance of UAV operators, and in a broader sense this ability affects performance on tasks ranging from driving to flight training. Regarding collaboration with automated systems, it has been suggested that operators with lower attentional control rely more heavily on the automation.

Spatial ability is another relevant factor for human control of automation. Amongst others, spatial ability is essential for navigation, UAV control, visual scanning and target detection (Chen & Barnes, 2014).

Gaming experience has been investigated as a predictor of human trust in HAT systems. This makes sense because many computer games resemble realistic HAT scenarios. In strategy games, for example, gamers often control groups of elements (like tanks) that are implemented via AI. In general, gaming experience has been found to increase trust in autonomous systems. In some cases, this leads to over-trust, which can be alleviated using simple interventions like presenting negative quotes about system performance from previous operators (Clare et al., 2015).

Besides self-efficacy being an outcome from a team interaction, it is also a trait that humans bring into the interaction. In working with automation, research shows that self-efficacy can have an influence on complacency and thereby on performance. Specifically, having low self-efficacy can induce strategies that increase the likelihood of automation-induced complacency. In other words, people who think poorly of themselves are more likely to follow the system. People with high self-efficacy, on the other hand, are more likely to offload too little work to the system, resulting in high workload (Prinzel, 2002).

Prior experience with robots leads to more positive attitudes towards robots and increased trust (Takayama, Takayama, & Pantofaru, 2016). Prior experience is mentioned here as a human characteristic but might also be fostered through interaction and can therefore also be seen as a mediator.

2.3.2 Robot characteristics

According to a meta-analysis by Hancock et al. (2011), characteristics of autonomous systems that contribute to trust in the system fall into either of two categories. The first category consists of performance-based factors, related to the reliability and effectiveness of the robot. The second category is related to robot attributes such as 'personality', and visual appearance. Both categories contribute to the development and maintenance of trust in the robotic system. Performance-based factors include failure rate (Hancock et al., 2011; Merritt & Ilgen, 2008), reliability, and false alarm rate (Hancock et al., 2011).

Anthropomorphism, the degree to which an agent exhibits human characteristics, has a major impact on how the agent is perceived and, consequently, the schema's that are triggered in interaction. Research indicates, for example, that emotions are easily attributed even to robots that show simple dog behaviour (Gacsi et al., 2016). It has been found that anthropomorphic agents are trusted more than machines (Visser et al. 2016).

Appearance also has an effect on feeling responsible for the task. One study showed, for example, that participants relied more on a human-like than a machine-like collaborator (Hinds et al., 2004). In a study by Biswas and Murray (2017) it was found that the preference for humanlike machines even extended to biased performance as humans preferred to interact with a robot containing biased algorithms as compared to unbiased robots.

A central problem of many advanced autonomous systems is that they (partially) rely on algorithms that remain opaque to the user, which has stimulated research into the topic of explainable AI (XAI, Lent, Rey, Fisher & Mancuso, 2004). The goal of this area of research is to develop intuitive measures and explanations that systems can use to calibrate operator's mental models and trust (e.g., Waa, Diggelen, & Neerincx, n.d.). Furthermore, we know from social psychology that a lack of justification from a teammate can negatively affect interaction. In all, even though the technical challenges here are immense, progress on XAI is likely essential to further develop HAT effectiveness.

The design of AI can greatly influence the way information is communicated and thus the way it interacts with its teammates. Some research has been done in the field of aviation involving the effects of speech vs text as a modality of communications. It was found that using speech instead of textual communication had an advantage, e.g. noticing more deviations in the monitoring task (Stedmon et al., 2007). The same research describes another experiment where the effects of real versus synthesised speech were compared. It was found that by using real speech there was an increased level of trust in the system and reduced reaction times (Stedmon et al., 2007). In general, the use of speech is of great influence on the HAT and trust development, particularly as speech invokes characteristics resembling a personality.

2.3.3 Gaps related to input

The number of human and robot characteristics is quite substantial and they all have their own effects on the outcomes and mediators previously mentioned. With regard to human characteristics it is particularly noteworthy that effects may change over time, for example due to positive or negative experiences in interacting with AI. To date, however, experiments mostly addressed short term effects and longitudinal data are scarce. With regard to robot design more attention should be paid to input variables affecting human-robot dialogue. In the context of explainable AI more insight is needed on how robot design and type of reasoning affect human trust and motivation. Human-like AI triggers mental processes that may be positive for social interaction but less advantage for task performance. In designing AI both factors, and their mutual effects, need to be taken into account.

All the input variables mentioned above could also matter in the interaction between Lieutenant Bob and Smart Bud: does Bob for example feel in control in debating with AI and does he understand the arguments provided? How does Bob assess the level of accuracy of Smart Bud, does he have any positive or negative experiences? And would he decide differently if Smart Bud was a robot rather than an algorithm informing him through his personal device?

3.0 DISCUSSION

3.1 Human appraisal

Most studies on HATs focused on the question of how to optimise task processes as to maximise task-related outcomes. As noted above, one reason for this emphasis might be that AI is often considered as technological advancement further increasing effectiveness and speed of current operations. However, even in kinetic operations, team performance is not solely driven by cognitive mechanisms such as having adequate SA or relevant task knowledge. Going back to the scenario we started with, lieutenant Bob might well feel compromised by Smart Buds superior situation awareness or restrained in his autonomy by constantly being monitored by Smart Bud. To date, knowledge is lacking concerning the influence of such motivational factors, i.e. human appraisal, in cooperating with AI, and how these factors interact with task-related factors. We expect human appraisal to become even more important in the future however, as AI becomes more and more engrained in team functioning. A view incorporating a wider view on motivational factors would significantly elaborate research questions with regard to HAT. For example, if Smart Bud is able to detect my level of workload or fatigue, how would that affect my level of experienced autonomy, and would I need emotional support from Smart Bud when in a dangerous situation? Furthermore, we know from research on perceived autonomy that it predicts persistence and adherence, and improves effective performance especially in complex tasks (Deci & Ryan, 2008).

Research on human appraisal in HATs should take results from the social-robotics field into account, as this field particularly focused on such factors. Note, however, that these studies did not consider human-automation interaction in a team or in a military context, but instead investigated AI as human support or assistant (for instance being a buddy to elderly, Looije, Neerincx & Cnossen,, 2010; Wada, Shibata, Musha, & Kimura, 2008). A main gap is therefore research on human appraisal in team interaction, which dominates the human team literature, but is rather absent in the HAT literature.

Another important outcome, or property, of sociotechnical systems is resilience. Resilience, or “the capability of recovering safely and efficiently from abnormal events” (Patriarca, Bergström, Di Gravio, & Costantino, 2018) has recently been reviewed in the context of HATs (Matthews, Barber, Teo, Wohleber, & Lin, 2016), and is related to many of the mediators and inputs we discuss here. For the military context however, we believe that ‘mental resilience’, or the individual ability to cope with events that can have an impact on mental well-being, should also be investigated in the light of HATs. This is because we know that team properties can affect mental resilience of individual members (e.g., Kamphuis, van Hemert, van Wouwe, van den Berg, & van Boxmeer, 2012), and replacing humans by systems strongly affects those properties.

3.2 Human information processing

As far as mediators are concerned many studies deal with adequate (shared) mental models as a prerequisite for optimal team performance. Team members have expectations with regard to each other’s knowledge level, reliability, needs and behaviour with regard to joint task performance. Human aware AI, for example, seeks to design accurate user models allowing the AI to best anticipate on human needs and support, as to optimally complement human task performance.

However, as also noted by Patterson (2017), military thinking as well as research on human-autonomy teaming usually assumes, implicitly, a rational and deliberate decision-making process (Williams, 2010). For instance, military theory often uses the OODA loop (observe, orient, decide, act) to model decision-making. In reality however, many of these steps are taken unconsciously for most of our decisions. A useful model that describes human decision-making is Kahneman’s dual process theory. It proposes that humans make decisions using one of two systems. System 1 is fast, intuitive, automatic and heuristic-based, while system 2 is slow, effortful and analytic (Kahneman, 2011). Williams (2010) discusses the relevance of decision-making theory for military operations and presents several examples of biases occurring in practice. Even though such biases will also occur in HATs, empirical research on both presence and mitigation is still sparse (see also Bosch and Bronkhorst, 2018).

Some research suggests that humans prefer biased robots over unbiased ones as they are more humanlike (Biswas 2017). Mirnig et al. (2017) showed, for example, that robots who gave erroneous instructions during a LEGO task, were perceived as more likeable, but equally intelligent, than robot who gave correct instructions. Salem et al. (2013) showed the same effect in a moving task where the robot gave instructions on where to put the items. The robot in the incongruent speech-gesture condition (e.g. say it goes left and point right) was perceived more likeable than the robot in the congruent speech-gesture condition. Moreover, they found that participants were more positive about future interactions in the incongruent condition. Participants did perform significantly worse, however, when they got incongruent instructions.

Presumably, intuitive processes will become more important when level of agency of AI increases. Not only because of increased intentionality ascribed to AI, but also because of all kinds of social processes. These unconscious and automatic processes have substantial influences on human thinking, task performance and social interactions. Social robotics use these natural tendencies through anthropomorphised AI. By making AI more human-like or animal-like, specific cognitive schema's are automatically triggered, supporting goals to be attained. These goals are obviously different than goals in military operations. However, also in a military context unconscious processes play significant roles, which should be taken into consideration when designing AI.

3.3 Conclusion

AI will become more and more important in future military operations. We concluded from our literature review that current research in the military domain mostly focuses on how to optimize task outcomes, and that AI is still mainly viewed as an 'add-on' to human decision makers (Spiegeleire, et al., 2017). However, as also discussed by Bosch & Bronkhorst (2018) there is a trend towards collaborative decision making, where AI is regarded as a true team-player. Such a situation would require profound mutual awareness, in order to jointly adapt to complex and dynamic situations. In addition to this, we argued that increased agency of AI will also trigger social behaviours, as these are deeply anchored into our (largely) social brain. So in addition to cognitive challenges, more insight is needed into how humans react socially to artificial teammates to further optimize HATs. As these aspects are well recognised in other domains such as health care (see for example Looije, Neerincx, & Cnossen, 2010), a main challenge would be to combine these different research strands, and to investigate how trade-offs are made with regard to task and relation in mixed human-AI teams.

REFERENCES

[1] Baker, A. L., & Keebler, J. R. (2017). Advances in Human Factors in Robots and Unmanned Systems, *499*, 331–340. https://doi.org/10.1007/978-3-319-41959-6

[2] Bosch, K., van den & Bronkhorst, A. (2018). Human-AI Cooperation to Benefit Military Decision Making. *Human-AI Cooperation to Benefit Military Decision Making*, *STO-MP-IST*, 1–12.

[3] Boyd, J. (1995). The Essence of Winning and Losing, by John R. Boyd, (August).

[4] Chen, J. Y. C., & Barnes, M. J. (2014). Human–Agent Teaming for Multirobot Control: A Review of Human Factors Issues. *IEEE Transactions on Human-Machine Systems*, *44*(1), 13–29. https://doi.org/10.1109/THMS.2013.2293535

[5] Clare, A. S., Cummings, M. L., & Repenning, N. P. (2015). Influencing trust for human-automation collaborative scheduling of multiple unmanned vehicles. *Human Factors*, *57*(7), 1208–1218. https://doi.org/10.1177/0018720815587803

[6] Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology/Psychologie Canadienne*, *49*(1), 14–23. https://doi.org/10.1037/0708-5591.49.1.14

[7] Ekelhof, M. A. C. (2015). Autonome Wapens. *Militaire Spectator*, *184*(5), 232–245.

[8] Gao, F., Clare, A. S., Macbeth, J. C., & Cummings, M. L. (2013). Modeling the Impact of Operator

Trust on Performance in Multiple Robot Control,. *AAAI Spring Symposium: Trust and Autonomous Systems*, 16–22.

[9] Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2017). When Will AI Exceed Human Performance? Evidence from AI Experts, 1–21.

[10] Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in Organizations: From Input-Process-Output Models to IMOI Models. *Annual Review of Psychology*, *56*(1), 517–543. https://doi.org/10.1146/annurev.psych.56.091103.070250

[11] Kamphuis, W., van Hemert, D. A., van Wouwe, N., van den Berg, C. E., & van Boxmeer, L. E. L. M. (2012). Een model van mentale veerkracht Hoe kan Defensie herstel na uitzendingen bevorderen ? *Militaire Spectator*, *181*(11), 495–507.

[12] Lent, M. Van, Rey, M., Fisher, W., & Mancuso, M. (2004). An Explainable Artificial Intelligence System for Small-unit Tactical Behavior. *In Proceedings of the Sixteenth Conference on Innovative Applications of Artificial Intelligence*, 900–907.

[13] Looije, R., Neerincx, M. A., & Cnossen, F. (2010). Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human Computer Studies*, *68*(6), 386–397. https://doi.org/10.1016/j.ijhcs.2009.08.007

[14] Mathieu, J. E., Heffner, T. S., Goodwin, G., Salas, E., & Cannon-Bowers, J. (2000). The Influence of Shared Mental Models on Team Process and Performance. *Journal of Applied Psychology*, *85*(2), 273–283.

[15] Nyholm, S. (2017). Attributing Agency to Automated Systems: Reflections on Human–Robot Collaborations and Responsibility-Loci. Science and engineering ethics, 1-19.

[16] Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors*, *56*(3), 476–488. https://doi.org/10.1177/0018720813501549

[17] Prinzel, L. J. I. (2002). The Relationship of Self-Efficacy and Complacency in Pilot-Automation Interaction. *Technical Memorandum No. TM-2002-211925, NASA.*

[18] Salas, E., Sims, D. E., & Shawn Burke, C. (2005). Is there A "big five" in teamwork? *Small Group Research*, *36*(5), 555–599. https://doi.org/10.1177/1046496405277134

[19] Sanders, T. L., Macarthur, K., Volante, W., Hancock, G., Macgillivray, T., Shugars, W., & Hancock, P. A. (2017). Trust and Prior Experience in Human-Robot Interaction, (2011), 1809–1813. https://doi.org/10.1177/1541931213601934

[20] Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation. *Human Factors*, *58*(3), 377–400. https://doi.org/10.1177/0018720816634228

[21] Spiegeleire, S., de Maas, M., & Sweijs, T. (2017). *Artificial Intelligence and the Future of Defense – Strategic Implications for Small- and Medium-Sized Force Providers*. *The Hague Centre for Strategic Studies (HCSS)*. Retrieved from http://www.hcss.nl/report/artificial-intelligence-and-future-defense

[22] Stanton, N. A., Salmon, P. M., Walker, G. H., Salas, E., & Hancock, P. A. (2017). State-of-science:

situation awareness in individuals, teams and systems. *Ergonomics*, *60*(4), 449–466. https://doi.org/10.1080/00140139.2017.1278796

[23] Stowers, K., Oglesby, J., Sonesh, S., Leyva, K., Iwig, C., & Salas, E. (2017). A framework to guide the assessment of human-machine systems. *Human Factors*, *59*(2), 172–188. https://doi.org/10.1177/0018720817695077

[24] Takayama, L., Takayama, L., & Pantofaru, C. (2016). Influences on Proxemic Behaviors in Human-Robot Interaction Influences on Proxemic Behaviors in Human-Robot Interaction, (NOVEMBER 2009), 5495–5502. https://doi.org/10.1109/IROS.2009.5354145

[25] Visser, E. J., de Pak, R., & Shaw, T. H. (2018). From “automation” to “autonomy”: The importance of trust repair in human-machine interaction. *Ergonomics*, *139*, 1–33. https://doi.org/10.1080/00140139.2018.1457725

[26] Waa, J., van der Diggelen, J., & Neerincx, M. (n.d.). The design and validation of an intuitive confindence measure. https://doi.org/10.1145/1235

[27] Wada, K., Shibata, T., Musha, T., & Kimura, S. (2008). Robot therapy for elders affected by dementia. *IEEE Engineering in Medicine and Biology Magazine*, *27*(4), 53–60. https://doi.org/10.1109/MEMB.2008.919496

[28] Wong, L., Kolditz, T. A., Millen, R. A., & Potter, T. M. (2003). Why they fight: combat motivation in the Iraq war. ARMY WAR COLL STRATEGIC STUDIES INST CARLISLE BARRACKS PA.

[29] You, S., & Robert, L. (2017). Teaming up with Robots : An IMOI (Inputs-Mediators-Outputs- Inputs) Framework of Human-Robot Teamwork, (October).