Human Factors and Systems Interaction, Vol. 52, 2022, 9–16
https://doi.org/10.54941/ahfe1002133

Exploring Trust in Unmanned Systems with the Maritime Unmanned System Trust Game

Francesca de Rosa and Christopher Strode

NATO STO Centre for Maritime Research and Experimentation, 19126 La Spezia, Italy

ABSTRACT

Maritime Unmanned Systems (MUS) are a quickly maturing technology, but for some security and defence operations we are still lacking the required doctrinal development. Trust in MUS has been identified as one of the main pillars of the adoption of such technology in the short term. Trust in automation and autonomy is an important and complex mental construct, which has been demonstrated to be based on several factors. With the goal of increasing the understanding of the operational tasks in future environments and steering future scientific developments, we designed the Maritime Unmanned Systems Trust (MUST) Game. The MUST Game is an analytical game which captures beliefs, attitude and perspectives of the participants with respect to the employment of MUS in maritime applications. This game aims at better understanding the relation between trust factors and MUS in missions. The game explored how players make decisions with respect to MUS deployments as the scenario threat level increases. This allows the capturing the important information on the trade-offs related to MUS use having an impact on mission planning activities (e.g., endurance, logistics, maintenance, cost, number of assets, security and type of assets). This paper presents the results of the analysis of the data collected through the deployment of the MUST Game in three distributed exercises.

Keywords: Autonomous systems, Trust, Analytical game, Knowledge acquisition

INTRODUCTION

MUS can be defined as an "unmanned system operating in the maritime environment (subsurface, surface, air) whose primary component is at least one unmanned vehicle. An unmanned vehicle is a [...] vehicle that does not carry a human operator and can: a. be operated autonomously or remotely; b. be expendable or recoverable; and c. can carry lethal or non-lethal payloads" (e.g., sensors) (Evangelio et al., 2012). Maritime Unmanned Systems (MUS) are moving towards higher levels of technological maturity. However, it appears that within the domain of maritime operations the doctrinal development that would allow their use is still lacking (Le Bourhis et al., 2018). A recent assessment on the status of policies and doctrines related to the use of MUS for maritime operations (Le Bourhis et al., 2018) identifies building trust in MUS as a potential game changer. Trust in automation and autonomy is a complex mental construct, which goes beyond mere familiarity and exposure to a certain technology. A comprehensive approach is needed to ensure an appropriate understanding of this construct and its operationalisation. Many studies have focused on technology related aspects of MUS (e.g., (Bradley et al., 2019)), while in this study we focus rather on the human assessment regarding MUS. The aim is to better understand foreseen benefits, potential issues in technology uptake by operators and which factors are important in the decisions to deploy manned, mixed or fully autonomous forces. This information are expected to inform the next generation of decision support systems available to the warfighters and maritime operators. With the goal of better understanding the relation between trust and MUS in ASW, contributing to the human-system integration efforts in this domain, the authors have developed the Maritime Unmanned Systems Trust (MUST) Game (de Rosa et al., 2022). This is an analytical wargame that aims at capturing beliefs, attitude and perspectives of the participants with respect to the employment of MUS. The adoption of MUS could prove to be effective (e.g., enhancement of situational awareness, reduction of human workload and enhancement of the operational performance) and would complement the manned platforms, positively impacting persistence, versatility, survivability, risk reduction and cost reductions. However, many are the elements that decision makers will have to trade-off at different levels (e.g. strategic, operational and tactical). Therefore, while focusing on trust in MUS, the MUST Game ensures the collection of relevant information also in relation to other mission planning factors, such as endurance, covertness, logistics, maintenance and cost. Moreover, the proposed operational analysis approach seeks to employ gaming methods to explore the real world complexities of large scale deployments of unmanned assets in maritime missions. The remainder of this paper is organised as follows: the Trust Related Work Section provides a brief overview of the concept of trust in current scientific developments; the MUST Game Section introduces the analytical game employed and related work; the Results Section presents the outcome of the analysis of the game data; finally, the Conclusion Section summarises the lessons learned and future work.

TRUST RELATED WORK

Trust has been demonstrated to be an important mediating factor with respect to the employment of technologies, especially in high-risk situations (Groom, 2007). For example, low levels of trust might affect the willingness of humans to rely on information and suggestions provided by the technology. On the contrary, inappropriately high levels of trust might lead to overreliance and misuse (Hancock et al., 2011). Trust is an important factor that determines the adoption and use of new systems in operations. The three-factor model of human-robot trust (Hancock et al., 2011) has identified several underpinning factors (i.e., antecedents of trust) that pertain to three main categories and further analysed how they influence trust development. Specifically, such factors can be human-related, system-related or environmental. Human-related factors refer to either human characteristics (i.e., demographics, personality traits, attitude towards the systems, comfort with the system, self-confidence and propensity to trust) or ability-based aspects (i.e, attention capacity or engagement, expertise, competency, user workload, prior experience and user situational awareness). System-related factors include performance-based and attribute-based elements. The performance-based ones are system behavious, dependability, reliability, predictability, level of automation, failure rate, false alarm rate and transparency. Instead, attribute-based elements include proximity or co-location of system and user, system personality, adaptability, system type and anthropomorphism. Finally, environmental factors are divided in team collaboration factors (i.e., in-group membership, culture, communication, shared mental models) and tasking (i.e., task type, task complexity, multi-tasking requirements and the physical environment). In our study we refer to this model to drive the assessments around trust.

THE MUST GAME

The MUST Game is an analytical wargame based on the Knowledge Acquisition Analytical Game (K2AG) approach (de Rosa, 2020). This game combines the strengths of K2AG with elements typical of Disruptive Technology Assessment Games (DTAG) (Collins, 2014), such as cards containing information regarding new technologies. DTAGs are table-top seminar games aimed at assessing potential disruptive technologies described on Ideas of System (IoS) cards. This method has proven to be an efficient tool when assessing early stage prototypes or technologies that have not yet been used in military operations (Collins, 2014). K2AGs, instead, are games used as a knowledge acquisition tool. Knowledge acquisition embraces the extraction, structuring and organisation of expert knowledge to be encoded in intelligent systems. Successful K2AGs, include the Reliability Game (de Rosa, 2018; de Rosa, 2019) and the MARISA Game (de Rosa, 2020), where data has been used to design algorithms to be employed in intelligent systems on the basis of cognitive mimetic principles. K2AGs have proven to be very efficient and effective in terms of knowledge elicitation (i.e. time reduction, experiment simplicity and ability to extract the required qualitative and quantitative knowledge). Two main gaming aspects that characterise K2AGs are the use of knowledge cards (KCs) to render information and meta-information to the players and the use of innovative data gathering methods to easily collect players' beliefs. These elements are at the core of the MUST Game mechanics as well, which is enriched with a technology led confrontation on the basis of IoSs. Specifically, the MUST Game is an analytical wargame, aiming at collecting useful information to support the development of future Concept of Operations and the continuous development of decision support tools that will allow decision makers to fully exploit the use of MUS within maritime missions. The MUST Game has been deployed both as a manual and as a distributed wargame within a series of Table Top Exercises (TTXs). In fact, the pandemic crisis impacted the game design as well as the game deployment. Consequently, several game sessions have been played over a time frame of eight months, when circumstances allowed for it. The players are a combination of military and civilians with a background in maritime operations and autonomy. During the game relevant data on players' beliefs and decisions was collected. Moreover, we have been able to observe how participants with different backgrounds and expertise perceive the use of MUS in maritime missions under an evolving threat level. Furthermore, all the participants as part of the pre-game data collection performed a cart sorting mini-game focused on the perceived relevance to maritime operations of trust factors.

RESULTS

The cart sorting mini-game data resulted in an ordinal ranking of the trust factor from the most important to the least. The analysis of the collected data allowed the derivation of an aggregated ranking of the trust factors, though an ordinal consensus ranking problem approach. This consists in a decision-making problem where decision-makers are requested to rank a set of alternatives (or attributes) with regard to a set of criteria or one overall criterion. In our problem formulation the alternatives are the trust factors, while the criterion is the importance with respect to the development of trust in MUS for maritime missions. We computed the collective preference vector, the aggregated preference vector of the military participants and the aggregated preference vector of the civilian participants. These last two were derived to evaluate if substantial differences could be identified between the opinions of these sub-groups. It appears that when it comes to trust in MUS, system reliability and system failure rates are the most important factors in all the consensus rankings. Overall system-related performance-based factors (i.e., system reliability, system failure rate, system false alarms, system behaviour, system predictability and system dependability) were globally rated as the more import factors, while the system- related attribute-based factors (i.e, system adaptability, system type and system and user co-location) were considered less important. The results on system-related performance based antecedents are overall in line with the results of the human-autonomy interaction (HAI) and human-robot interaction (HRI) meta-analyses on trust antecedents (Hancock et al., 2011; Schaefer et al., 2016). These studies summarise how system related factors, particularly those that are performance based, have the largest influence on perceived trust. Human related ability based factors such as user situational awareness, user prior experience and user competency were included in the top ten factors, while those of user workload and user attention capacity were regarded as less important. Also environmental factors such as task type, multi-tasking requirements and physical environment were considered less important. In fact, the only environmental factor included in the ten most important factors is task complexity. However, we should notice the gap between the opinions provided by the military sub-group and the civilian subject matter experts. In fact, we observed how the civilian consensus vector includes in the top ten factors seven out of the eight system related performance based factors (i.e., system reliability, system failure rate, system false alarms, system behaviour, system predictability, system transparency and system dependability). The other three factors are human-related (i.e., user situational awareness, user workload, user prior experience). The military consensus vector, instead, exhibits a greater inclusion of human related factors (i.e, user competency, user situational awareness, user expertise and user prior experience) and environmental related ones (i.e., task complexity and multi-tasking requirements). Research on HRI and HAI (Hancock et al., 2011; Schaefer et al., 2016) has shown how human-related and environmental antecedents indeed have an effect on trust development. However, it also underlines the lack of adequate experimental support, as most studies in the past decades were technology focused. In the military consensus vector the user workload was rated very low. This is in direct contrast to the results obtained for the civilian consensus vector, which highly rated the importance of the user workload. This is an important observation as reduced user workload has actually been demonstrated to be correlated to the phenomenon known as out-of-the-loop performance and changes in individual intervention performances that might be introduced by automation and autonomy (Endsley et al., 2017). This discrepancy should be further explored in order to ensure that the workload factor is indeed negligible due to a low risk of out-of-the-loop performance in maritime missions involving the use of MUS. Both sub-groups agree on a lower ranking of the system-related attribute based factors, such as system-user co-location, system adaptability and system type. These factors should be further investigated. The aim of this further analysis would be to ensure that all the important factors that contribute to a successful accomplishment of the operational goals (i.e. technology factors, human-related factors and human-system integration factors) are correctly addressed in future research and development (R&D) projects as well as in the definition of the Concept of Use and Concept of Operations of MUS, harmonizing the different perspective between the scientific and the operational community along the doctrine, organization, training, materiel, leadership and education, personnel, facilities and interoperability (DOTMLPFI) dimensions. While it is not surprising that system reliability, system failure rates and system false alarms score within the top five factors that underpin trust in MUS, an interesting result relates to the ranking of factors such as system predictability and system transparency, which are often mentioned as key elements to build trust. In fact, we observed how the civilian consensus vector rates it as the third and tenth most important factors respectively, while the military consensus vector rates them considerably lower (i.e., thirteenth and fourteenth respectively).

Overall the complexity of the decision space and of the mental model of the players was observed. Examples of parameters that explicitly played an important role include: the military posture (i.e., shape, deter, defend), the tension level (i.e., peacetime, crisis, war), the desired posture (i.e., covert or overt assets), type of mission (e.g., hold at risk or protected passage), the complexity of the task (including required area coverage), the role of the player (e.g. an operational commander, politician or someone responsible for logistics), performances of available assets (e.g., speed, detection range, endurance), survivability, logistics and sustainability considerations (including launch and recovery from sea, air and land), a vulnerability and consequence assessment (i.e., prioritise the assets to be defended) and the patterns of life (POLs). In fact, deployments often result from a perceived change in the adversary POLs. Another important factor is reaction time, in fact the current capabilities do not allow to quickly deploy MUS in highly dynamic environments. This stresses the importance of the role of "augmentation" of the manned assets played by MUS and the need to carefully consider the complementarities when planning the deployment. In fact, we observed a tendency towards the deployment of mixed forces (manned and unmanned). The decision space factors are considerable and correlated. Moreover, they do not directly correspond to current mathematical optimisation models. However, a careful analysis of such factors could lead to a mapping of these different factors. Therefore, the results of this analysis suggest that while working on the refinement of the mathematical optimisation models to be included in advanced decision support systems for the use of MUS, attention should be payed to the development of an additional computational layer, able to bridge these models with the operators' mental ones. This could substantially increase the support provided to different kinds of users in different contexts and situations.

The data collected during gameplay included the self-reported frequency of the selection of factors impacting the decisions during the unfolding game scenario. The set of factors from which the players could select, included the trust factors, endurance, security, covertness and deployment time. During the scenario-based simulation the decision factors playing a major role appear to be endurance and user situational awareness. We observed how other relevant factors are the physical environment, covertness, deployment time, system type and user workload, while the ones playing a smaller role are system failure rates, user competency, user attentional capacity, user prior experience, system predictability, system-user co-location and system transparency. The military and civilian sub-groups agree on the factors with the lowest selection rate (e.g., system transparency, user prior experience and system predictability), while there is a substantial different percentage of selection with respect to the following factors, which are consistently selected more by the civilians: multi-tasking requirements, system adaptability, system false alarms, covertness, deployment time and user situational awareness. Military participants self-reported to consider in most cases endurance, user situational awareness, the physical environment and deployment time. Civilian participants, instead, self-reported to consider endurance, user situational awareness, covertness and deployment time in most cases. The covertness aspects strongly relate to political component. In general, the use of MUS is perceived as less provocative due to their covertness. The human component has been stressed as an important aspect to consider from different perspectives. In fact, it has been mentioned how further steps need to be taken to make this systems completely operational as these systems are far from being human ready. In fact, usability and human-system integration aspects have not been adequately tackled so far.

CONCLUSION

This paper presents the outcome of three distributed TTXs on the use of Maritime Unmanned Systems for maritime operations. The TTXs made use of the Maritime Unmanned System Trust (MUST) Game. Through the use of the MUST Game we were able to observe the relative importance of the trust factors and those factors affecting decision-making (i.e., cost, variety and number). The results appear to shine an interesting light on the perceived importance of the trust factors, which could have relevant implications on directing the future R&D efforts. Moreover, important observations emerged during gameplay relating to: (i) potential barriers and the political component of the use of MUS, (ii) the importance of the human component, (iii) the tendency towards employing mixed (manned and unmanned) forces and (iv) the need for decision support systems. Future work should investigate these aspects. Finally, the collected data will be further modelled to inform the design of dedicated decision support tools.

ACKNOWLEDGMENT

This research received support by NATO Allied Command Transformation (NATO-ACT) through the Autonomous Anti Submarine Warfare (A-ASW) programme of work of the NATO STO Centre for Maritime Research and Experimentation (CMRE).

REFERENCES

A. Evangelio, O. Nyaas, G. Yuzichuck, S. Sweeney, M. Zaragoz, M. Coffman, and J. Fox. NATO Guidance for developing Maritiem Unmanned Systems (MUS) capability. North Atlantic Treaty Organisation, Combined Joint Operations from the Sea Centre of Excellence, 2012.

Collins S. Disruptive Technology Assessment Game Handbook. North Atlantic Treaty Organisation, North Atlantic Treaty Organisation, Allied Command Transformation, 2014.

F. de Rosa and A. De Gloria. An analytical game for knowledge acquisition for maritime behavioral analysis systems. Applied Science, 10(2):591, 2020.

F. de Rosa, A. De Gloria, and A.-L. Jousselme. Analytical games for knowledge engineering of expert systems in support to situational awareness: the reliability game case study. Expert Systems with Applications, 138:112800–112811, 2019.

F. de Rosa, A.-L. Jousselme, and A. De Gloria. A Reliability Game for Source Factors and Situational Awareness Experimentation. International Journal of Serious Games, 5(2):45–64, 2018.

F. de Rosa, Strode, C. Games to support disruptive technology adoption: the MUST Game use case. International Journal of Serious Games, (under review)

F. de Rosa. Knowledge Acquisition Analytical Games: games for cognitive system design. PhD thesis, University of Genoa, Italy, 2020.

G. Le Bourhis, L. Bertonati, and M. Murphy. Maritime Unmanned Systems in ASW. Collaborative ASW study. North Atlantic Treaty Organisation, Combined Joint Operations from the Sea Centre of Excellence, 2018.

K. E. Schaefer, J. Y. C. Chen, J. L. Szalma, and P. A. Hancock. A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. Human Factors, 58(3):377–400, 2016.

M. Bradley, D. C. Tarraf, T. C. Whitmore, J. DeWeese, C. Kenney, J. Schmid, and P. DeLuca. Advancing Autonomous Systems - An Analysis of Current and Future Technology for Unmanned Maritime Vehicles. RAND Corporation, Santa Monica, California, 2019.

M. R. Endsley. From Here to Autonomy: Lessons Learned From Human-Automation Research. Human Factors, 59(1):5–27, 2017.

P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. de Visser, and R. Parasuraman. A meta-analysis of factors affecting trust in human-robot interaction. Human Factors, 53(5):517–527, 2011.