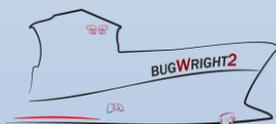


# **Human-Robot Teams.**

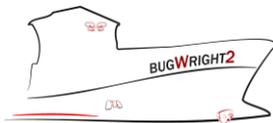
Spotlight on Psychological Acceptance Factors  
exemplified within the BUGWRIGHT2 Project

Nathalie Schauffel, Jan Gründling, Benjamin Ewerz,  
Benjamin Weyers, Thomas Ellwart





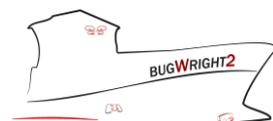
Research conducted by the units *Business Psychology* and *Human-Computer Interaction* of Trier University



Research conducted within the BUGWRIGHT2 project  
*Autonomous Robotic Inspection and Maintenance on Ship  
Hulls* ([www.bugwright2.eu](http://www.bugwright2.eu))



Research funded by the European Union's Horizon 2020  
research and innovation programme under Grant  
Agreement No. 871260



### **E-book Citation**

Schauffel, N., Gründling, J., Ewerz, B., Weyers, B., & Ellwart, T. (2022). Human-Robot Teams. Spotlight on Psychological Acceptance Factors exemplified within the BUGWRIGHT2 Project. *PsychArchives*.  
<https://doi.org/10.23668/psycharchives.5584>

This work is licensed under a Creative Commons Attribution 4.0 International License.  
<https://creativecommons.org/licenses/by/4.0/>

### **Acknowledgment**

We thank Helena Schroeder and our research assistants Lena M. Kaufmann, Maria-Elena Zavrakidou, Corinna Pfortner, Laura Peci, and Julia Lellmann for their scientific and administrative support.

### **Funding**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 871260.

### **Authors Contact**

**Nathalie Schauffel, M.Sc. \***

**Benjamin Ewerz, M.Sc.**

**Prof. Dr. Thomas Ellwart**

Trier University, Business Psychology

Universitätsring 15

54296 Trier, Germany

**Jan Gründling, M.Sc.**

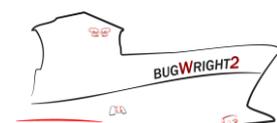
**Prof. Dr. Benjamin Weyers**

Trier University, Human-Computer Interaction

Behringstr. 21

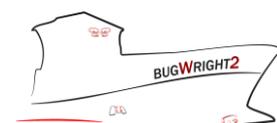
54296 Trier, Germany

\* Correspondence: Nathalie Schauffel, email: [schauffel@uni-trier.de](mailto:schauffel@uni-trier.de)



## Table of Contents

Abstract.....	1
Introduction .....	2
Factsheet 1: Acceptance of Human-Robot Interaction .....	4
Factsheet 2: Agent Transparency in a Human-Robot Context .....	7
Factsheet 3: Artificial Intelligence – Human-centered and Explainable .....	10
Factsheet 4: Attitudes towards Remote Inspection Technologies (RIT) .....	13
Factsheet 5: Basic Human Needs in Human-Robot Interaction.....	16
Factsheet 6: Cognitive Load (Theory) and Workload.....	19
Factsheet 7: Cybersickness – Feeling moved without being moved .....	22
Factsheet 8: Immersion and Presence in Human-Robot Interaction .....	25
Factsheet 9: Measuring Task Performance in Human-Robot Teams .....	28
Factsheet 10: Promoting Appropriate Trust in HRI through Design.....	31
Factsheet 11: Self-Efficacy in the Context of Human-Robot Interaction.....	35
Factsheet 12: Situation Awareness – Knowing what’s going on.....	38
Factsheet 13: SMART Work via SMART Work Design.....	41
Factsheet 14: Trust in Automation in a Multi-Robot System .....	44
Conclusion .....	48

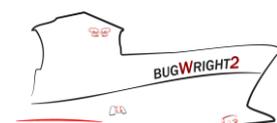


## Abstract

The project BUGWRIGHT2 is an interdisciplinary collaborative project co-funded by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 871260. The project aims to propel the digital-maritime revolution by developing an adaptable autonomous robotic solution for vessel-structure inspection and maintenance.

From a psychological point of view, the implementation of autonomous robots in the work context changes work tasks, roles, and responsibilities from an all-human to a human-robot team setting. Therefore, concepts of psychological team research, humane work design, and technology acceptance need to be considered to realize the full potential of robotic solutions in the maritime sector.

This e-book intends to spotlight 14 psychological topics identified as essential for the acceptance of an autonomous robotic solution developed within the BUGWRIGHT2 project through literature research and expert interviews. Each psychological topic is presented in a factsheet that summarizes the scientific input, provides appropriate literature recommendations, and concludes with recommendations for the BUGWRIGHT2 project. The factsheets are valuable for any researcher or practitioner interested or involved in implementing robotic solutions in a work environment.



## Introduction

The EU project *BUGWRIGHT2 (Horizon 2020) Autonomous Robotic Inspection and Maintenance on Ship Hulls and Storage Tanks* is an interdisciplinary EU research and innovation project that aims to develop an adaptable autonomous robotic solution for the inspection and maintenance of ship hulls and storage tanks. BUGWRIGHT2 intends to combine heterogeneous robotic technologies (in air, underwater, above water), virtual reality (VR), and augmented reality (AR), to contribute to revolutionizing ship inspections and maintenance.<sup>1</sup>

Transforming a currently mainly manual process including dry-dock, scaffolding, or human divers into a future robotic-supported process has far-reaching positive impacts regarding human safety (e.g., reduced risk of diving accidents or scaffolding collapses), the environment (e.g., reduced fuel consumption and pollution of the seas), and economy (e.g., reduced dry-dock times that result in lower inspection costs and more time-effective shipping).<sup>2</sup>

From a work psychology point of view, however, the implementation of adaptable autonomous robots also leads to fundamental changes of the underlying work process, roles, and responsibilities: The workflow as well as the related and required competencies change as single tasks become obsolete while new tasks might be relevant elsewhere. The team composition shifts from all-human teams to a human-robot team setting. The interaction between humans and robots might be supported by VR and AR. Responsibilities and liabilities have to be redefined for routines and emergencies. These changes bring psychological opportunities, but also risks and challenges. To reveal the full potential of human-robot interaction and minimize potential risks, psychologically successful human-robot collaboration requires the optimal interplay and calibration of human, technical, and organizational subsystems involved in a given work task (Karlton et al., 2017).

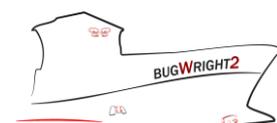
As (semi-)autonomous operations of inspection technologies relate to self-government and self-directed behavior of a robotic solution (O'Neill et al., 2020), the humans involved depend on the robot's performance, provided data, and decisions (e.g., the robotic system signals no critical damage on the outer ship hull). This interdependency, among others, requires well-calibrated trust in the robotic solution (i.e., avoidance of mistrust and overtrust), reasonable cognitive load, high situation awareness, and user acceptance. This guarantees short-term and long-term well-being of humans involved, in the sense of humane work design.

This e-book spotlights 14 psychological topics relevant to the success of human-robot teams in the field that relate to the human, technology, and organizational subsystem of human-robot teams. Reviewing psychological literature on the topics of human-robot teams (e.g., You & Robert, 2017), humane work design (e.g., Klonek & Parker, 2021), and technology acceptance (e.g., Bröhl et al., 2019; Venkatesh et al., 2016) the following factsheets provide a concise summary of the scientific state-of-the-art including the theoretical background and empirical relations. Furthermore, concrete recommendations for the BUGWRIGHT2 project are derived regarding the monitoring and promotion of these critical psychological factors.

---

<sup>1</sup> Official project page: [www.bugwright2.eu](http://www.bugwright2.eu)

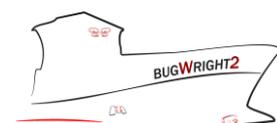
<sup>2</sup> Taken from the Description of the Action, Annex 1 to the Grant Agreement 871260.



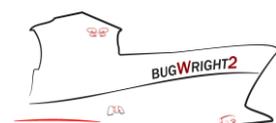
At a glance, this e-book presents how *agent transparency* and *explainable artificial intelligence (XAI)* contribute to increasing robot trust and acceptance. In addition, we focus on concepts closely related to human attention, such as *situation awareness*, that reflect the knowledge of the current circumstances or *cognitive load*, a concept rooted in learning science that can also be applied to problem-solving within human-robot team tasks. Two factsheets deal with phenomena in the context of virtual environments, namely *cybersickness* and *immersion and presence*. Different methods to measure *task performance* and key concepts of *technology acceptance* are also reviewed. Regarding cognitive-motivational factors, two factsheets focus on the topic of *trust* in human-robot teams. The impact of competence self-perceptions in human-robot teams is reviewed in the factsheet on *self-efficacy* which describes one's self-perceived confidence to succeed in a situation or task. The relevance of human *attitudes* in human-robot teams and possible methods of attitude change are presented. Regarding humane work design, we present the concept of *smart work design* as a valuable framework to analyze and evaluate human-robot work settings. Furthermore, we outline the critical role of *basic human needs* satisfaction in human-robot teams. The factsheets are presented in alphabetical order.

## References

- Bröhl, C., Nelles, J., Brandl, C., Mertens, A., & Nitsch, V. (2019). Human-robot collaboration acceptance model: Development and comparison for Germany, Japan, China and the USA. *International Journal of Social Robotics*, 11(5), 709–726. <https://doi.org/10.1007/s12369-019-00593-0>
- Karlton, A., Karlton, J., Berglund, M., & Eklund, J. (2017). HTO - A complementary ergonomics approach. *Applied Ergonomics*, 59(Pt A), 182–190. <https://doi.org/10.1016/j.apergo.2016.08.024>
- Klonek, F., & Parker, S. K. (2021). Designing SMART teamwork: How work design can boost performance in virtual teams. *Organizational Dynamics*, 50(1), 100841. <https://doi.org/10.1016/j.orgdyn.2021.100841>
- O'Neill, T., McNeese, N., Barron, A., Schelble, B., & O'Neill, T. (2020). Human-autonomy teaming: A review and analysis of the empirical literature. *Human Factors*, 18720820960865. <https://doi.org/10.1177/0018720820960865>
- Venkatesh, V., Thong, J., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, 17(5), 328–376. <https://doi.org/10.17705/1jais.00428>
- You, S., & Robert, L. P. (2017). Teaming up with robots: An IMOI (inputs-mediators-outputs-inputs) framework of human-robot teamwork. *International Journal of Robotic Engineering*, 2(1), 1–7. <https://doi.org/10.35840/2631-5106/4103>



## Factsheet 1: Acceptance of Human-Robot Interaction



## Factsheet 1: Acceptance of Human-Robot Interaction

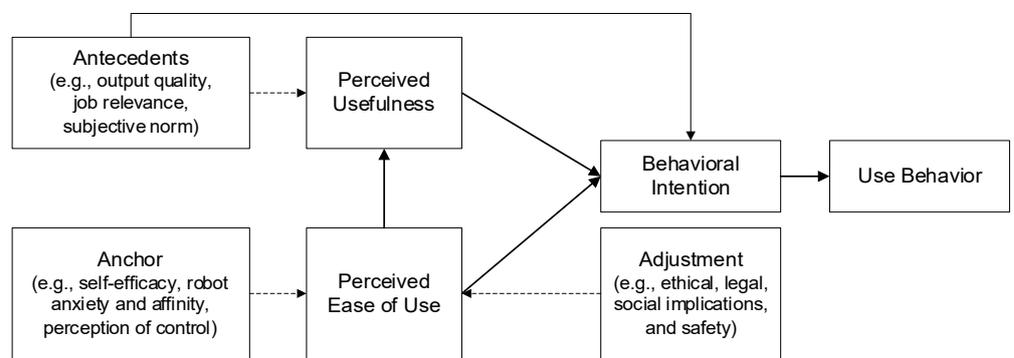
**Keywords** Technology acceptance; human-robot interaction; human-robot collaboration

**At a Glance** Well-established models of technology acceptance systematize technology-, person-, and context-related factors that determine the acceptance and use of robot technology in the field. The perceived usefulness and ease of human-robot interaction (HRI) are key determinants of future robot use.

### Scientific Input

**Theoretical Background** Multiple models of technology acceptance exist. The models systematize different types of user evaluations which cause a behavioral intention to use a system or actual use behavior of a system (Venkatesh et al., 2003). Rooted in the acceptance of information systems, technology acceptance models have been continuously extended (see Venkatesh et al., 2016) and transferred to HRI (Bröhl et al., 2019).

**Empirical Relations** Behavioral Intention (BI) describes the intention to use or to recommend a system to others. BI is a widely used proxy for actual system use (Venkatesh et al., 2003). The perceived usefulness and ease of use of a robot system are key determinants of BI and, thus, robot use (see Figure 1). Manifold antecedents impact the perceived usefulness of HRI but also have a direct effect on BI. These more traditional technology acceptance factors include the output quality, the robot’s relevance for the given work task, or existing behavioral norms (e.g., support of robot use). Person-related factors like self-efficacy (i.e., confidence to accomplish HRI), robot anxiety, and affinity or control perceptions impact “human decision-making processes” (Bröhl et al., 2019, p. 710) and thus the perceived ease of use as anchor variables. The two-sidedness of robots in the field both as an opportunity (e.g., to support humans) and a threat (e.g., fear of job loss, see Smids et al., 2019) further requests to include ELSI factors in models of HRI acceptance (Bröhl et al., 2019). ELSI stands for ethical, legal, and social implications (Bröhl et al., 2019). Researchers also integrated the topic of trust into technology acceptance (e.g., Belanche et al., 2012) and embedded core variables of technology acceptance against the background of basic human needs (e.g., Fathali & Okada, 2018).

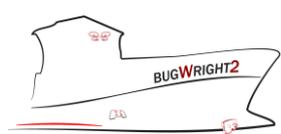


**Figure 1.** Model of human-robot interaction acceptance (modified, based on Bröhl et al., 2019).

### Recommendations for BUGWRIGHT2

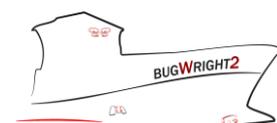
**Measurement, Promotion, Intervention, Etc.** Qualitative interviews supported the relevance of manifold acceptance factors within the specific application case of BUGWRIGHT2. Regarding person-related anchor variables, human perception of control and competence (i.e., self-efficacy) turned out to be particularly important within BUGWRIGHT2. Another factor vital for the acceptance of remote inspection technologies is trust. To measure individual factors of HRI acceptance, well-established self-rating scales provide starting points for customized evaluation (see recommended literature).

**Recommended Literature** Bröhl, C., Nelles, J., Brandl, C., Mertens, A., & Nitsch, V. (2019). Human-robot collaboration acceptance model: Development and comparison for Germany, Japan, China and the USA. *International Journal of Social Robotics*, 11(5), 709–726. <https://doi.org/10.1007/s12369-019-00593-0>  
 Venkatesh, V., Thong, J., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, 17(5), 328–376. <https://doi.org/10.17705/1jais.00428>

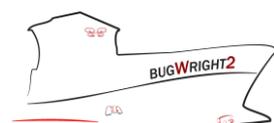


## References

- Belanche, D., Casaló, L. V., & Flavián, C. (2012). Integrating trust and personal values into the technology acceptance model: The case of e-government services adoption. *Cuadernos De Economía Y Dirección De La Empresa*, 15(4), 192–204. <https://doi.org/10.1016/j.cede.2012.04.004>
- Bröhl, C., Nelles, J., Brandl, C., Mertens, A., & Nitsch, V. (2019). Human–robot collaboration acceptance model: Development and comparison for Germany, Japan, China and the USA. *International Journal of Social Robotics*, 11(5), 709–726. <https://doi.org/10.1007/s12369-019-00593-0>
- Fathali, S., & Okada, T. (2018). Technology acceptance model in technology-enhanced OCLL contexts: A self-determination theory approach. *Australasian Journal of Educational Technology*, 34(4). <https://doi.org/10.14742/ajet.3629>
- Smids, J., Nyholm, S., & Berkers, H. (2019). Robots in the workplace: A threat to—or opportunity for—meaningful work? *Philosophy & Technology*, 17(9–10), 608. <https://doi.org/10.1007/s13347-019-00377-4>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
- Venkatesh, V., Thong, J., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, 17(5), 328–376. <https://doi.org/10.17705/1jais.00428>

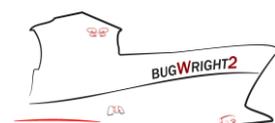


## Factsheet 2: Agent Transparency in a Human-Robot Context



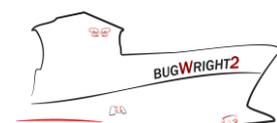
## Factsheet 2: Agent Transparency in a Human-Robot Context

<b>Keywords</b>	Human-robot interaction; transparency; interface design; autonomy
<b>At a Glance</b>	In a human-robot team, transparency of the actions, plans, and reasoning process of an autonomous robot helps the human to better understand the robot, which in turn leads to higher trust and overall performance in human-robot teams.
<b>Scientific Input</b>	
<b>Definition and Theoretical Background</b>	Interactions between robots or autonomous agents and humans will get more complex as systems increase in their autonomy (Lyons, 2013). Therefore, the necessity increases for the human operator to accurately understand the agent's actions, reasoning process, and projections in the actual situation. Making this information accessible to the user is a concept known as agent transparency (Selkowitz, 2017). Agent transparency can be defined as the "quality of an interface pertaining to its abilities to afford an operator's comprehension about an intelligent agent's intention, performance, future plans, and reasoning process" (Chen et al., 2014). According to the Situation Awareness-Based Agent Transparency model, there are three levels of information an agent needs to convey to maintain a transparent interaction with the human: the agent's current status/actions (level 1), the agent's reasoning process behind the actions (level 2) and the agent's projections/predictions (level 3). The purpose of transparency is to facilitate interaction in a human-robot team.
<b>Empirical Relations</b>	Results from several studies suggest that high agent transparency is associated with increased operator performance, trust, situation awareness, perceived usability, and understanding of the agent (e.g., Chen et al., 2016; Harbers et al., 2011; Mercado et al., 2016). In contrast, low agent transparency results in decreased performance and higher human complacency (Mercado et al., 2016; Wright et al., 2016). O'Neill et al. (2020) pointed out that transparency plays an interactive role in the effects of an agent's reliability, as human operators showed higher trust and performance when they were aware of the lower reliability (higher transparency) of the agent. However, it is important to consider the type and amount of information to provide agent transparency. For example, Wright et al. (2016) found that information transparency in the form of projected outcomes and uncertainty information has hindered operator performance. Selkowitz et al. (2016) indicated that providing the operator with information in the high transparency condition improved the operator's situation awareness more than in the very high transparency condition. Thus, the highest transparency level did not always produce the best outcome (Bhaskara, 2017). In summary, the general trends of results suggest that transparency has a positive impact on various outcome variables (e.g., Bhaskara, 2017; O'Neill et al., 2020). However, this does not seem to be a linear relationship as too much transparency can have a negative impact on performance (Bhaskara, 2017). To sum up, the concepts of performance (see Factsheet 9), situation awareness (see Factsheet 12), and agent transparency are closely related to each other.
<b>Recommendations for BUGWRIGHT2</b>	
<b>Interface Design</b>	Agent transparency is an important variable for interface design. Implementing information about the agent's actions, reasoning process, and projections in the user interface supports the operator's ability to develop an accurate understanding of the environment and the autonomous agent. In addition, the perceived usability increases with higher levels of transparency (Chen et al., 2016). In contrast, a poorly designed display could conceal the meaning behind the information and render the information useless (Selkowitz et al., 2017).
<b>Recommended Literature</b>	Lyons, J. B. (2013). Being transparent about transparency: A model for human-robot interaction. <i>2013 AAAI Spring Symposium Series</i> Mercado, J. E., Rupp, M. A., Chen, J. Y., Barnes, M. J., Barber, D., & Procci, K. (2016). Intelligent agent transparency in human-agent teaming for Multi-UxV management. <i>Human factors</i> , 58(3), 401-415. <a href="https://doi.org/10.1177/0018720815621206">https://doi.org/10.1177/0018720815621206</a>

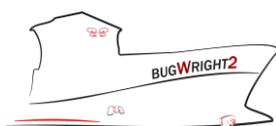


## References

- Bhaskara, A. (2017). The operationalisation of agent transparency and evidence for its impact on key human-autonomy teaming variables.
- Chen, J. Y., Barnes, M. J., Selkowitz, A. R., Stowers, K., Lakhmani, S. G., & Kasdaglis, N. (2016). Human-Autonomy Teaming and Agent Transparency. In J. Nichols, J. Mahmud, J. O'Donovan, C. Conati, & M. Zancanaro (Eds.), *Companion Publication of the 21st International Conference on Intelligent User Interfaces* (pp. 28–31). ACM. <https://doi.org/10.1145/2876456.2879479>
- Chen, J. Y., Procci, K., Boyce, M., Wright, J., Garcia, A., & Barnes, M. (2014). Situation awareness-based agent transparency. Technical Report: ARL-TR-6905, Aberdeen Providing Ground. <https://doi.org/10.1117/12.2263194>
- Harbers, M., Bradshaw, J. M., Johnson, M., Feltovich, P., van den Bosch, K., & Meyer, J.-J. (2011). Explanation and coordination in human-agent teams: A study in the BW4T Testbed. In *2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology* (pp. 17–20). IEEE. <https://doi.org/10.1109/WI-IAT.2011.83>
- Lyons, J. B. (2013). Being transparent about transparency: A model for human-robot interaction. *2013 AAAI Spring Symposium Series*.
- Mercado, J. E., Rupp, M. A., Chen, J. Y., Barnes, M. J., Barber, D., & Procci, K. (2016). Intelligent agent transparency in human-agent teaming for Multi-UxV management. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 58(3), 401-415. <https://doi.org/10.1177/0018720815621206>
- O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2020). Human-autonomy teaming: A review and analysis of the empirical literature. *Human Factors*. <https://doi.org/10.1177/0018720820960865>
- Selkowitz, A. R., Lakhmani, S. G., Larios, C. N., & Chen, J. Y. (2016). Agent transparency and the autonomous squad member. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 1319-1323. <https://doi.org/10.1177/1541931213601305>
- Selkowitz, A. R., Larios, C. A., Lakhmani, S. G., & Chen, J. Y. (2017). Displaying Information to Support Transparency for Autonomous Platforms. In P. Savage-Knepshield & J. Chen (Eds.), *Advances in Intelligent Systems and Computing. Advances in Human Factors in Robots and Unmanned Systems* (Vol. 499, pp. 161–173). Springer International Publishing. [https://doi.org/10.1007/978-3-319-41959-6\\_14](https://doi.org/10.1007/978-3-319-41959-6_14)
- Wright, J. L., Chen, J. Y., Barnes, M. J., & Hancock, P. A. (2016). Agent reasoning transparency's effect on operator workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 249-253. <https://doi.org/10.1177/1541931213601057>

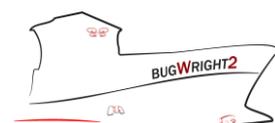


## **Factsheet 3: Artificial Intelligence – Human-centered and Explainable**



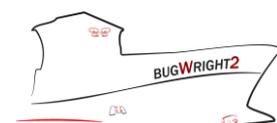
### Factsheet 3: Artificial Intelligence – Human-centered and Explainable

<b>Keywords</b>	Explainable artificial intelligence; artificial intelligence; trustworthiness; machine learning; deep learning; human-centered AI
<b>At a Glance</b>	Explainable Artificial Intelligence (XAI) aims at making the results of AI and its inner working intelligible for humans (i.e., mental model of AI). Especially post-hoc explanations of the AI (e.g., machine learning) are vital in complex systems to increase system trust and acceptance.
<b>Scientific Input</b>	
<b>Definition and Systematization</b>	Research on XAI contributes to human-centered AI, aiming to make it easier for future generations to understand, trust in, and manage AI (Adadi & Berrada, 2018) and, thus, optimize the perceived interaction, safety, fairness, and informativity, among others, when interacting with an AI system. Research on XAI aims to make the results of AI systems more intelligible for humans (Adadi & Berrada, 2018) and to make their inner workings clearer (Barredo Arrieta et al., 2020) by explaining them. These explanations can be broadly classified into intrinsic (i.e., by AI design) or post-hoc (i.e., afterward) methods of XAI. In BUGWRIGHT2, especially post-hoc methods are vital. In such a complex system, explainability does not fully result from the AI design (e.g., if the model is too complex) but is provided after AI use. Post-hoc methods are diverse and based on features, examples, rules, concepts, and strategies underlying machine learning (Molnar, 2019).
<b>Empirical Relations</b>	Accurate mental models (i.e., mental representation of knowledge) are vital for well-calibrated trust (i.e., not too high, or too low). XAI helps to establish accurate mental models of AI, as humans tend to erroneously transfer mental models of human functioning to the functioning of AI (e.g., emotion detection). Here, empirical research shows that XAI (e.g., visualizing the area on which an AI decision is based) can diminish this misleading tendency (Heimerl et al., 2020). Further, research by Mertes et al. (2020) indicates that different example-based explanations (i.e., counterfactuals) positively influence trust, satisfaction, and accurate decision-making. Weitz et al. (2019; 2020) show that the more human an AI seems, the more trustworthy it is perceived.
<b>Recommendations for BUGWRIGHT2</b>	
<b>1. Provide Post-Hoc Explanations</b> <b>2. Verify the Mental Models</b>	Existing research on XAI is young but clearly points out two things: 1. Explanations will influence end-users perceptions and attitudes towards the BUGWRIGHT2 system: Especially regarding autonomous detection of corrosion and excessive biofouling but also regarding automated mission planning, post-hoc explanations of the AI system help increase system trust and acceptance by the end-users. Here, multiple methods already exist (see Adadi & Berrada, 2018). 2. Not every explanation is a helpful explanation: XAI should help to establish accurate mental models. Ask end-users what they think about the AI functioning, the decision-making process, and the results. Verifying the accuracy of AI mental models helps to establish a well-calibrated level of trust and to avoid counteracting effects of overtrust or mistrust.
<b>Recommended Literature</b>	Weitz, K., Schiller, D., Schlagowski, R., Huber, T., & André, E. (2019). "Do you trust me?" Increasing user-trust by integrating virtual agents in explainable AI interaction design. <i>Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents</i> (pp. 7-9). <a href="https://doi.org/10.1145/3308532.3329441">https://doi.org/10.1145/3308532.3329441</a> Weitz, K., Schiller, D., Schlagowski, R., Huber, T., & André, E. (2020). "Let me explain!": Exploring the potential of virtual agents in explainable AI interaction design. <i>Journal on Multimodal User Interfaces</i> , 1-12. <a href="https://doi.org/10.1007/s12193-020-00332-0">https://doi.org/10.1007/s12193-020-00332-0</a>

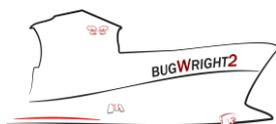


## References

- Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58(3), 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Heimerl, A., Weitz, K., Baur, T., & Andre, E. (2020). Unraveling ML models of emotion with NOVA: Multi-level explainable AI for non-experts. *IEEE Transactions on Affective Computing*, 1. <https://doi.org/10.1109/TAFFC.2020.3043603>
- Mertes, S., Huber, T., Weitz, K., Heimerl, A., & André, E. (2020). *This is not the texture you are looking for! Introducing novel counterfactual explanations for non-experts using generative adversarial learning.*
- Molnar, C. (2019). *Interpretable machine learning: A guide for making Black Box Models interpretable.* Lulu.
- Weitz, K., Schiller, D., Schlagowski, R., Huber, T., & André, E. (2019). "Do you trust me?" Increasing user-trust by integrating virtual agents in explainable AI interaction design. *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents* (pp. 7-9). <https://doi.org/10.1145/3308532.3329441>
- Weitz, K., Schiller, D., Schlagowski, R., Huber, T., & André, E. (2020). "Let me explain!": Exploring the potential of virtual agents in explainable AI interaction design. *Journal on Multimodal User Interfaces*, 1-12. <https://doi.org/10.1007/s12193-020-00332-0>

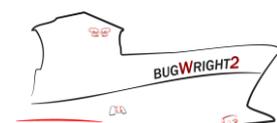


## **Factsheet 4: Attitudes towards Remote Inspection Technologies (RIT)**



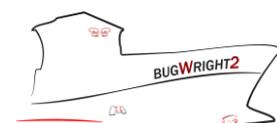
## Factsheet 4: Attitudes towards Remote Inspection Technologies (RIT)

<b>Keywords</b>	Attitudes towards robots; robot affinity; interest; remote inspection technologies
<b>At a Glance</b>	A positive attitude towards robots is vital for effective human-robot teamwork. Therefore, cognitive, affective, and conative components of attitudes are relevant and can be measured via self-reports. Direct experience, imagined interaction, and knowledge transfer are methods to influence attitudes.
<b>Scientific Input</b>	
<b>Definition</b>	An attitude describes a person's relative endurance of subjective positive or negative tendency towards an attitude object (e.g., new technology, robot; cf., Eagly & Chaiken, 1993).
<b>Theoretical Background: ABC-model</b>	Drawing on the three-component model of attitudes by Rosenberg and colleagues (1960), attitudes comprise a cognitive (i.e., knowledge), an affective (i.e., emotional), and a conative (i.e., behavioral) component. The three components are interrelated.
<b>Empirical Relations</b>	Attitudes play a central role in models of technology acceptance stating that attitudes cause a behavioral intention (see Venkatesh et al., 2003). Attitudes, therefore, are used to predict consumer and user behavior, even though routines or other situational influences can prevent that a behavior intention leads to actual behavior (Sheeran, 2002). Qualitative research shows that techno-related attitudes are diverse, ranging from enthusiasm to skepticism (Kerschner & Ehlers, 2016). Attitudes toward robots as person-related user characteristics impact trust in service robots (Miller et al., 2021). Trust is an essential core requirement for effective human-robot interaction.
<b>Recommendations for BUGWRIGHT2</b>	
<b>Attitude Assessment</b>	Attitudes become observable through the definition and measurement of indicators. To measure attitude towards robots or RITs, multiple indicators can exist. Apart from psychophysiological indicators (e.g., pulse rate) or observable features (e.g., facial expressions, gestures), in practical applications, subjective experiences (e.g., self-rating questionnaires) are mainly used to measure attitudes. Here, validated measures exist to quantify an individual's tendency regarding technology interaction (e.g., Franke et al., 2019), attitudes towards robots (e.g., Carpinella et al., 2017; Ninomiya et al., 2015; Nomura et al., 2008; Robert, 2021) or technophobia (e.g., Khasawneh, 2018; Sinkovics et al., 2002).
<b>Methods of Attitude Change</b>	Multiple evidence-based methods of attitude change exist that can be transferred to the context of HRI, and specifically RITs: <i>User participation</i> = The involvement of users in the robot design process lowers robot anxiety and increases positive robot attitude (Reich-Stiebert et al., 2019) <i>Imagined context</i> = The imagination of a first positive interaction with RITs. Thoughts and associations during the imagination are documented on a worksheet. <i>Context dependence</i> = Reflection on RITs in different work tasks. Open discussion on human-robot interaction and RITs in each task (e.g., What constitutes HRI in this domain?, What features/strengths/weaknesses do RITs have in this task?) <i>Knowledge transfer and persuasion</i> = Communication of empirical evidence on the topic of HRI and RITs, the effects of robot implementation on work processes and job profiles, and the impact on economy, safety, environment. Realistic information on robot abilities but also uniquely human abilities.
<b>Recommended Literature</b>	Reich-Stiebert, N., Eyssel, F., & Hohnemann, C. (2019). Involve the user! Changing attitudes toward robots by user participation in a robot prototyping process. <i>Computers in Human Behavior</i> , 91, 290–296. <a href="https://doi.org/10.1016/j.chb.2018.09.041">https://doi.org/10.1016/j.chb.2018.09.041</a> Robert, L. P. (2021, July). A measurement of attitude toward working with robots (AWRO): A compare and contrast study of AWRO with negative attitude toward robots (NARS). <i>International Conference on Human-Computer Interaction</i> , 12763, 288–299. Springer. <a href="https://doi.org/10.1007/978-3-030-78465-2_22">https://doi.org/10.1007/978-3-030-78465-2_22</a>

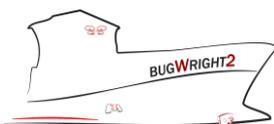


## References

- Carpinella, C. M., Wyman, A. B., Perez, M. A., & Stroessner, S. J. (2017). The robotic social attributes scale (rosas). In B. Mutlu, M. Tscheligi, A. Weiss, & J. E. Young (Eds.), *Hri'17: proceedings of the acm/ieee international conference on human-robot interaction: March 6-9, 2017, Vienna, Austria* (pp. 254–262). IEEE. <https://doi.org/10.1145/2909824.3020208>
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.
- Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ati) scale. *International Journal of Human-Computer Interaction*, 35(6), 456–467. <https://doi.org/10.1080/10447318.2018.1456150>
- Kerschner, C., & Ehlers, M.-H. (2016). A framework of attitudes towards technology in theory and practice. *Ecological Economics*, 126, 139–151. <https://doi.org/10.1016/j.ecolecon.2016.02.010>
- Khasawneh, O. Y. (2018). Technophobia: Examining its hidden factors and defining it. *Technology in Society*, 54, 93–100. <https://doi.org/10.1016/j.techsoc.2018.03.008>
- Miller, L., Kraus, J., Babel, F., & Baumann, M. (2021). More than a feeling—interrelation of trust layers in human-robot interaction and the role of user dispositions and state anxiety. *Frontiers in Psychology*, 12, 592711. <https://doi.org/10.3389/fpsyg.2021.592711>
- Ninomiya, T., Fujita, A., Suzuki, D., & Umemuro, H. (2015). Development of the multi-dimensional robot attitude scale: Constructs of people's attitudes towards domestic robots. . In: Tapus A., André E., Martin JC., Ferland F., Ammi M. (eds) Social Robotics. ICSR 2015. Lecture Notes in Computer Science, vol 9388. Springer, Cham. [https://doi.org/10.1007/978-3-319-25554-5\\_48](https://doi.org/10.1007/978-3-319-25554-5_48)
- Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2008). Prediction of human behavior in human—robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Transactions on Robotics*, 24(2), 442–451. <https://doi.org/10.1109/TRO.2007.914004>
- Reich-Stiebert, N., Eyssel, F., & Hohnemann, C. (2019). Involve the user! Changing attitudes toward robots by user participation in a robot prototyping process. *Computers in Human Behavior*, 91, 290–296. <https://doi.org/10.1016/j.chb.2018.09.041>
- Robert, L. P. (2021, July). A measurement of attitude toward working with robots (AWRO): A compare and contrast study of AWRO with negative attitude toward robots (NARS). *International Conference on Human-Computer Interaction*, 12763, 288–299. Springer. [https://doi.org/10.1007/978-3-030-78465-2\\_22](https://doi.org/10.1007/978-3-030-78465-2_22)
- Rosenberg, M. J. (Ed.). (1960). *Attitude organization and change: An analysis of consistency among attitude components*. Yale University Press.
- Sheeran, P. (2002). Intention-behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, 12(1), 1–36. <https://doi.org/10.1080/14792772143000003>
- Sinkovics, R. R., Stöttinger, B., Schlegelmilch, B. B., & Ram, S. (2002). Reluctance to use technology-related products: Development of a technophobia scale. *Thunderbird International Business Review*, 44(4), 477–494. <https://doi.org/10.1002/tie.10033>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>

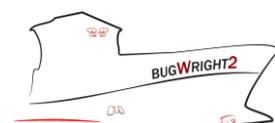


## Factsheet 5: Basic Human Needs in Human-Robot Interaction



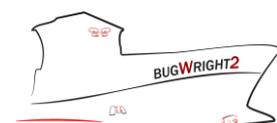
## Factsheet 5: Basic Human Needs in Human-Robot Interaction

<b>Keywords</b>	Self-determination theory; basic human needs; human-robot interaction
<b>At a Glance</b>	Robots at work can be perceived as an opportunity versus an offense for the fulfillment of basic human needs. Robots should be designed to support the need for competence, autonomy, relatedness, physical safety, and status.
<b>Scientific Input</b>	
<b>Definition and Theoretical Background</b>	According to the <i>Self-Determination Theory</i> (SDT, Deci & Ryan, 2000), there are three basic human needs (i.e., need for competence, relatedness, and autonomy) that are universal (i.e., everybody has them) and innate (i.e., not learned). Further, in the context of human-robot interaction, the need for physical safety (Pervez & Ryu, 2008) and status are important (Smids et al., 2019). Human needs "specify psychological nutrients that are essential for ongoing psychological growth, integrity, and well-being" (Deci & Ryan, 2000, p. 229). Their satisfaction is crucial for technology acceptance and use because people seek interaction that supports need satisfaction.
<b>Basic Human Needs and HRI</b>	Team behavior (i.e., supervisory or colleague) influences need satisfaction (Deci et al., 2017; Rynek et al., 2021). In hybrid teams consisting of humans and robots, robot behavior and features influence human need satisfaction and, thus, technology acceptance as well as human health.
<b>Need for Competence:</b> Desire to feel effective and experience masterful behavior	Competence self-perceptions ("I can") impact technology use, performance, and well-being (Brosnan, 1998; Igbaria & Iivari, 1995; Marsh et al., 2017). Training and user experience serve as central sources of positive competence self-perceptions. In addition, the robot's interaction style influences the user's subjective perception of competence. Zafari et al. (2019) showed that robots giving individual feedback to the user (e.g., encouraging, praising) results in higher human competence self-perception than robots using another interaction style (i.e., task-oriented, see Factsheet 11 on self-efficacy).
<b>Need for Relatedness:</b> Desire to be connected to others	Smids et al. (2019) showed that robots at work can be both an opportunity and an offense to the need for relatedness depending on the work task and robot type. If robots substitute repetitive tasks, the need for relatedness can be supported as additional time is left for interpersonal contact with other colleagues. In addition, robot interaction itself can be perceived as socially interactive. If robots replace valued co-workers, the need for relatedness can be frustrated.
<b>Need for Autonomy:</b> Desire to self-initiate and self-regulate one's own behavior	The robot's level of autonomy influences the operator's robot control strategies (Zhou et al., 2019). In the context of robotic decision-support, a trade-off between robot automation and user control exists (Rühr et al., 2019). If high robot autonomy takes decision-making power away from human end-users, the need for autonomy might be frustrated (Smids et al., 2019).
<b>Need for Physical Safety:</b> Desire to maintain physical health	Health (mental and physical) is considered the most basic human need that will be pursued by all individuals before any other need (Doyal & Gough, 1984). Consequently, the implementation of robots at work should sufficiently reduce existing safety risks, without creating new safety risks.
<b>Need for Status:</b> Desire to be valued by others	Employment gives people social and economic status (Super, 1980). Introducing robots at work changes job profiles and roles. To maintain or improve the perceived status, robots should not take over the tasks that are high in social recognition. Instead, new evolving job roles (i.e., robot supervision, robot calibration) should expand skills and lead to higher social recognition (Smids et al., 2019).
<b>Recommendations for BUGWRIGHT2</b>	
<b>ROOS: Robots as an Opportunity or an Offense to the Self</b>	For BUGWRIGHT2 a need-based approach for evaluating the human-robot interaction is helpful. In work analyses, current need-thwarting elements of the hull inspection process can be identified and optimized in the future automated work environment. Tracking basic human needs satisfaction may lead to high compliance and acceptance of new technology. Especially satisfaction of the needs for physical safety and autonomy are critical factors for the social acceptance of BUGWRIGHT2.
<b>Recommended Literature</b>	Smids, J., Nyholm, S., & Berkers, H. (2019). Robots in the workplace: A threat to—or opportunity for—meaningful work? <i>Philosophy &amp; Technology</i> , 33, 503–522. <a href="https://doi.org/10.1007/s13347-019-00377-4">https://doi.org/10.1007/s13347-019-00377-4</a>

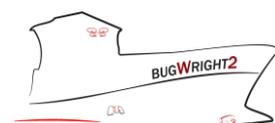


## References

- Brosnan, M. J. (1998). The impact of computer anxiety and self-efficacy upon performance. *Journal of Computer Assisted Learning*, 14(3), 223–234. <https://doi.org/10.1046/j.1365-2729.1998.143059.x>
- Deci, E. L., Olafsen, A. H., & Ryan, R. M. (2017). Self-determination theory in work organizations: The state of a science. *Annual Review of Organizational Psychology and Organizational Behavior*, 4(1), 19–43. <https://doi.org/10.1146/annurev-orgpsych-032516-113108>
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. [https://doi.org/10.1207/S15327965PLI1104\\_01](https://doi.org/10.1207/S15327965PLI1104_01)
- Doyal, L., & Gough, I. (1984). A theory of human needs. *Critical Social Policy*, 4(10), 6–38. <https://doi.org/10.1177/026101838400401002>
- Igbaria, M., & Iivari, J. (1995). The effects of self-efficacy on computer usage. *Omega the International Journal of Management Science*, 23(6), 587–605. [https://doi.org/10.1016/0305-0483\(95\)00035-6](https://doi.org/10.1016/0305-0483(95)00035-6)
- Marsh, H. W., Martin, A. J., Yeung, A. S., & Craven, R. G. (2017). Competence self-perceptions. In A. J. Elliot (Ed.), *Handbook of competence and motivation* (pp. 85–115). Guilford Press.
- Pervez, A., & Ryu, J. (2008). Safe physical human robot interaction – past, present and future. *Journal of Mechanical Science and Technology*, 22(3), 469–483. <https://doi.org/10.1007/s12206-007-1109-3>
- Rynek, M., Ellwart, T., Peiffer, H., Endres, E., & Moldzio, T. (2021). Threats to professional roles in part-time leadership. *Zeitschrift Für Arbeits- Und Organisationspsychologie A&O*, 35(5), Article 0932-4089/a000362, 166. <https://doi.org/10.1026/0932-4089/a000362>
- Rühr, A., Hess, T., & Berger, B. (2019). Can I control my robo advisor? Trade-offs in automation and user control in (Digital) investment management [Paper presentation]. Twenty-fifth Americas Conference on Information Systems, Cancun, Mexico.
- Smids, J., Nyholm, S., & Berkers, H. (2019). Robots in the workplace: A threat to—or opportunity for—meaningful work? *Philosophy & Technology*, 33, 503–522. <https://doi.org/10.1007/s13347-019-00377-4>
- Super, D. E. (1980). A life-span, life-space approach to career development. *Journal of Vocational Behavior*, 16(3), 282–298. [https://doi.org/10.1016/0001-8791\(80\)90056-1](https://doi.org/10.1016/0001-8791(80)90056-1)
- Zafari, S., Schwaninger, I., Hirschmanner, M., Schmidbauer, C., Weiss, A., & Koeszegi, S. T. (2019, October). “You are doing so great!” – The effect of a robot’s interaction style on self-efficacy in HRI. In *2019 28<sup>th</sup> IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (pp. 1–7). IEEE. <https://doi.org/10.1109/RO-MAN46459.2019.8956437>
- Zhou, J., Zhu, H., Kim, M., & Cummings, M. L. (2019). The impact of different levels of autonomy and training on operators’ drone control strategies. *ACM Transactions on Human-Robot Interaction*, 8(4), 1–15. <https://doi.org/10.1145/3344276>

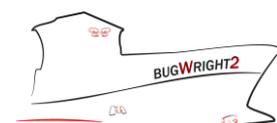


## Factsheet 6: Cognitive Load (Theory) and Workload



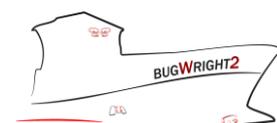
## Factsheet 6: Cognitive Load (Theory) and Workload

<b>Keywords</b>	Cognitive load; mental load; working memory
<b>At a Glance</b>	Cognitive Load is originally understood as the amount of resources from the working memory needed to complete a learning task. This can also be applied to problem-solving. The demand of working memory resources can be induced by different aspects such as task material, circumstances, and information processing (Sweller, 1988).
<b>Scientific Input</b>	
<b>Definition</b>	Cognitive Load is understood as the number of resources from the working memory needed to achieve task completion.
<b>Theoretical Background</b>	<p>Cognitive Load consists of the main components: intrinsic load, extraneous load and germane load (Sweller, 2010). According to Sweller (2011), they fulfill the following functions:</p> <p><i>Intrinsic load</i> means the resources required because of the immanent level of difficulty found in the material that is to be worked on (e.g., the degrees of freedom and thereby induced complexity in remote controlling a robotic system).</p> <p><i>Extraneous load</i> means the resources required because of the way information is presented to the subject or the actions that need to be taken to assess critical information for task completion (e.g., using a 3D representation for displaying positional data versus a numerical representation).</p> <p><i>Germane load</i> means the resources that are required to process and construct patterns and sort the perceived information into categories, which enables a person to transfer knowledge to the long-term memory or draw conclusions from perceived information.</p> <p>In the context of BUGWRIGHT2, it is not necessary to measure each component individually. It is rather needed to get an overall picture of the worker's workload. The construct (mental) workload can be understood as the result of cognitive load factors and individual characteristics (Galy, 2012).</p>
<b>Empirical Relations</b>	Mental Workload is the measured variable when it comes to assessing the cognitive demand that a worker experiences during a task. It is related to performance, fatigue (Fan & Smith, 2017), and worker satisfaction (Khandan et al., 2012). Moreover, it has been found that the use of AI technologies, as in the BUGWRIGHT2 project, can significantly reduce workload (Buettnner, 2013). Furthermore, there are attempts to integrate the real-time measurement of workload at work by physiological measures (Ramakrishnan et al., 2021) as well as by subjective measures (Mach et al., 2019). This aims to enable a possible artificial intelligence system to consider the "co-worker's" mental load and adapt its behavior to their level.
<b>Recommendations for BUGWRIGHT2</b>	
<b>Measurement</b>	<p>Cognitive Load measurement is primarily conducted by using means to assess the mental effort, which includes performance, physiological, and questionnaire measurements (Miller, 2001).</p> <p>Questionnaires:</p> <ul style="list-style-type: none"> <li>• NASA-TLX: 6 items + optional rating procedure (Hart, 2006; Sharek, 2011)</li> <li>• SWAT: 9 items (Luximon &amp; Goonetilleke, 2001; Rubio et al., 2004)</li> </ul> <p>Physiological measures:</p> <ul style="list-style-type: none"> <li>• Heart rate</li> <li>• Heart rate variability</li> <li>• Blood pressure</li> <li>• Pupil dilatation</li> <li>• Electroencephalogram (EEG)</li> </ul> <p>Performance measures:</p> <ul style="list-style-type: none"> <li>• Primary task performance</li> <li>• Secondary task performance</li> </ul>
<b>Recommended Literature</b>	Plass, J. L., Moreno, R., & Brünken, R. (Eds.). (2010). <i>Cognitive load theory</i> . Cambridge University Press. <a href="https://doi.org/10.1017/CBO9780511844744">https://doi.org/10.1017/CBO9780511844744</a>

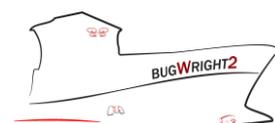


## References

- Brünken, R., Seufert, T., & Paas, F. (2010). Measuring Cognitive Load. In J. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory* (pp. 181-202). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511844744.011>
- Buettner, R. (2013). Cognitive workload of humans using artificial intelligence systems: Towards objective measurement applying eye-tracking technology. In *Lecture Notes in Computer Science. KI 2013: Advances in Artificial Intelligence* (Vol. 8077, pp. 37–48). Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-40942-4\\_4](https://doi.org/10.1007/978-3-642-40942-4_4)
- Fan, J., & Smith, A. P. (2017). The impact of workload and fatigue on performance. In *Communications in computer and information science. Human mental workload: Models and applications* (Vol. 726, pp. 90–105). Springer International Publishing. [https://doi.org/10.1007/978-3-319-61061-0\\_6](https://doi.org/10.1007/978-3-319-61061-0_6)
- Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 83(3), 269–275. <https://doi.org/10.1016/j.ijpsycho.2011.09.023>
- Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9), 904–908. <https://doi.org/10.1177/154193120605000909>
- Khandan, M., Roshan Zamir, S., & Maghsoudipour, M. (2012). Survey of workload and job satisfaction relationship in a productive company. *Iran Occupational Health Journal*, 9(1). <http://ioh.iums.ac.ir/article-1-706-en.html>
- Luximon, A., & Goonetilleke, R. S. (2001). Simplified subjective workload assessment technique. *Ergonomics*, 44(3), 229–243. <https://doi.org/10.1080/00140130010000901>
- Mach, S., Gründling, J. P., Schmalfuß, F., & Krems, J. F. (2019). How to assess mental workload quick and easy at work: A method comparison. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20<sup>th</sup> congress of the International Ergonomics Association (IEA 2018)*. Springer International Publishing. [https://doi.org/10.1007/978-3-319-96068-5\\_106](https://doi.org/10.1007/978-3-319-96068-5_106)
- Miller, S. (2001). *Literature Review. Workload measures*. National Advanced Driving Simulator. <http://www.nads-sc.uiowa.edu/publicationstorage/200501251347060.n01-006.pdf>
- Ramakrishnan, P., Balasingam, B., & Biondi, F. (2021). Cognitive load estimation for adaptive human-machine system automation. In *Learning Control* (pp. 35–58). Elsevier. <https://doi.org/10.1016/B978-0-12-822314-7.00007-9>
- Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology*, 53(1), 61–86. <https://doi.org/10.1111/j.1464-0597.2004.00161.x>
- Sharek, D. (2011). A useable, online NASA-TLX tool. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 55(1), 1375–1379. <https://doi.org/10.1177/1071181311551286>
- Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2), 257–285. [https://doi.org/10.1207/s15516709cog1202\\_4](https://doi.org/10.1207/s15516709cog1202_4)
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. Springer New York. <https://doi.org/10.1007/978-1-4419-8126-4>
- Sweller, J. (2010). Cognitive Load Theory: Recent Theoretical Advances. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 29–47). Cambridge University Press. <https://doi.org/10.1017/CBO9780511844744.004>

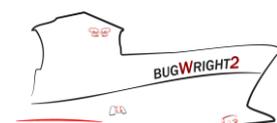


## **Factsheet 7: Cybersickness – Feeling moved without being moved**



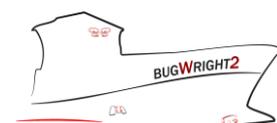
## Factsheet 7: Cybersickness – Feeling moved without being moved

<b>Keywords</b>	Cybersickness; virtual reality; motion sickness
<b>At a Glance</b>	A person might feel symptoms of sickness when they are moving in a virtual reality scene and not moving in reality. Because many people are affected by it, research has developed design guidelines to mitigate cybersickness.
<b>Scientific Input</b>	
<b>Definition</b>	Users of virtual reality, in which they seem to be moving in the virtual environment while remaining physically motionless, may experience cybersickness (Davis et al., 2014).
<b>Theoretical Background</b>	Since virtual reality's debut, cybersickness has been an issue. Nausea, disorientation, headaches, perspiration, and eye strain are just some of the symptoms that might accompany cybersickness. The exact origin of cybersickness is unknown and the physiological reactions that accompany it are also unspecified. Poison hypothesis, postural instability theory, and sensory conflict theory are the three most popular ideas for what causes cybersickness (Davis et al., 2014). It is estimated that more than 60% of first-time users of early VR headsets (i.e., from 1994-2010) experienced negative symptoms. Among these negative symptoms, motion sickness was predominant, dependent on criteria, equipment, and others. Approximately 5% did not experience any negative symptoms while other 5% quit earlier (Stanney et al., 2020).
<b>Empirical Relations</b>	In the BUGWRIGHT2 context, the most relevant finding might be that cybersickness correlates negatively with cognitive performance (Kim et al., 2005; Mittelstaedt et al., 2019; Nelson et al., 2000; Stanney et al., 2002). This suggests that the performance of affected VR users suffers from increased cybersickness. Moreover, cybersickness has severe negative effects on the user's wellbeing (van der Spek et al., 2010). In a very thorough literature review, Weech et al. (2019) conclude that cybersickness might be inversely correlated to presence.
<b>Recommendations for BUGWRIGHT2</b>	
<b>Measurement, Promotion, Intervention, Etc.</b>	<p>While cybersickness is a phenomenon quite specific to virtual reality, it should not be ignored in the scope of BUGWRIGHT2. As soon as any user interface integrates the slightest use of virtual reality, this becomes a very relevant topic. Therefore, a collection of measurement and mitigation techniques can be found below.</p> <p><i>Measurement.</i> Cybersickness can be subjectively measured by validated scales such as the Simulator Sickness Questionnaire, Nausea Profile, or Research and Brand's susceptibility survey (Davis et al., 2014).</p> <p><i>Intensification.</i> Cybersickness is intensified by lag, flicker, and wrong calibration. Participants with strong control in a virtual environment are shown to be less sensitive to cybersickness and can better predict future movements. Longer virtual reality exposure results in more bouts of cybersickness and symptom intensity, demanding longer adaptation times (Davis et al., 2014). Stanney et al. (2020) recommend the following techniques to mitigate cybersickness:</p> <ul style="list-style-type: none"> <li>• Minimize or remove motion parallax signals during first VR/AR exposure to allow consumers to adapt to the VR/AR experience. Motion parallax = Objects moving at a constant speed across the visual field appear to move further within one frame when closer to the user.</li> <li>• Stepwise increase of the motion parallax cues over time.</li> <li>• If a user becomes irritated, reduce the motion parallax cues once more.</li> <li>• Include options to reduce the amount of needed head movement/motion parallax.</li> <li>• Use teleportation (note: can increase disorientation and hinder spatial awareness).</li> <li>• Include concordant motion (e.g., motion-base to reduce visual-vestibular conflicts).</li> <li>• Request viewers to actively align their head/body with the virtual motion.</li> <li>• To minimize visual scene motion, limit or delay forward speed and acceleration.</li> </ul>
<b>Recommended Literature</b>	Davis, S., Nesbitt, K., & Nalivaiko, E. (2014). A systematic review of cybersickness. In <i>Proceedings of the 2014 Conference on Interactive Entertainment</i> . ACM. <a href="https://doi.org/10.1145/2677758.2677780">https://doi.org/10.1145/2677758.2677780</a>

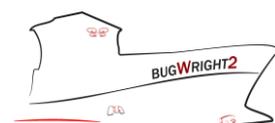


## References

- Davis, S., Nesbitt, K., & Nalivaiko, E. (2014). A systematic review of cybersickness. In *Proceedings of the 2014 Conference on Interactive Entertainment*. ACM. <https://doi.org/10.1145/2677758.2677780>
- Kim, Y. Y., Kim, H. J., Kim, E. N., Ko, H. D., & Kim, H. T. (2005). Characteristic changes in the physiological components of cybersickness. *Psychophysiology*, 42(5), 616–625. <https://doi.org/10.1111/j.1469-8986.2005.00349.x>
- Mittelstaedt, J. M., Wacker, J., & Stelling, D. (2019). VR aftereffect and the relation of cybersickness and cognitive performance. *Virtual Reality*, 23(2), 143–154. <https://doi.org/10.1007/s10055-018-0370-3>
- Nelson, W. T., Roe, M. M., Bolia, R. S., & Morley, R. M. (2000). *Assessing simulator sickness in a see-through HMD: Effects of time delay, time on task, and task complexity*. Fort Belvoir, VA. <https://doi.org/10.21236/ADA430344>
- Stanney, K. M., Kingdon, K. S., Graeber, D., & Kennedy, R. S. (2002). Human performance in immersive virtual environments: Effects of exposure duration, user control, and scene complexity. *Human Performance*, 15(4), 339–366. [https://doi.org/10.1207/S15327043HUP1504\\_03](https://doi.org/10.1207/S15327043HUP1504_03)
- Stanney, K. M., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., Weech, S., & Fulvio, J. M. (2020). Identifying causes of and solutions for cybersickness in immersive technology: Reformulation of a research and development agenda. *International Journal of Human-Computer Interaction*, 36(19), 1783–1803. <https://doi.org/10.1080/10447318.2020.1828535>
- van der Spek, E. D., Houtkamp, J. M., & Toet, A. (2010). Cybersickness influences the affective appraisal of a virtual environment. *The Open Virtual Reality Journal*, (2), 26–31. <https://edepot.wur.nl/191613>
- Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology*, 10, 158. <https://doi.org/10.3389/fpsyg.2019.00158>

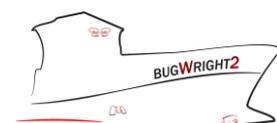


## Factsheet 8: Immersion and Presence in Human-Robot Interaction



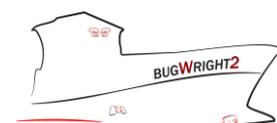
## Factsheet 8: Immersion and Presence in Human-Robot Interaction

<b>Keywords</b>	Immersion; presence; virtual reality; augmented reality
<b>At a Glance</b>	A person feels immersion when they are involved in and shift their attention towards the goals of virtual reality (VR) or augmented reality (AR). They feel like they are in the virtual world. Content and environment as well as technical factors influence immersion.
<b>Scientific Input</b>	
<b>Definition</b>	Immersion is a state of deep mental involvement in which the subject may experience disassociation from the awareness of the physical world due to a shift in their attentional state (Agrawal et al., 2020).
<b>Theoretical Background</b>	Agrawal et al. (2020) state that immersion can be divided into two different notions of the concept: psychological and perceptual immersion. <i>Psychological immersion</i> is understood as a user's psychological state when they are involved, absorbed, engaged, or engrossed. The attention is shifted towards the virtual world and its goals. <i>Perceptual immersion</i> is the state of being surrounded or experiencing multisensory stimulation whilst blocking/overpowering stimuli from the real environment. This sight aligns with what Slater (2009) describes. However, Agrawal's (2020) psychological immersion is very similar to a very broadly known concept called presence. Thus, in the following, we use the term immersion interchangeably with perceptual immersion and the term presence with psychological immersion. For BUGWRIGHT2, both concepts are relevant.
<b>Empirical Relations</b>	Immersion in the context of BUGWRIGHT2 should be understood as a variable that gives insight into how rich the channels are that a user interface uses for delivering information about the data collected by the robot system. For example, a VR interface where the user can travel around the whole ship and have a look at every detail should be considered as more immersive than a desktop interface (Slater et al., 2010). However, it is not always necessary or helpful to provide the highest possible amount of immersion but rather to find the most appropriate level for the specific use case (Bowman & McMahan, 2007). Therefore, immersion is part of a set of variables that measure the quality of the user interface in terms of enabling the user to conduct appropriate analytics on the collected data. As the perception of it is determined by the degree of immersion a system provides (Agrawal et al., 2020), presence is an important subjective variable to gain insight into how the user perceives the use of an immersive system. As a higher degree of presence is associated with a higher level of trust (Salanitri et al., 2016), it is also an essential variable in human-robot teams.
<b>Recommendations for BUGWRIGHT2</b>	
<b>Measurement, Promotion, Intervention, Etc.</b>	<p>Presence as the subjective construct can be measured by these techniques:</p> <ul style="list-style-type: none"> <li>• Independent Television Commission Sense of Presence Inventory (ITC-SOPI) (Lessiter et al., 2001) – 44 items</li> <li>• Measurement, Effects, Conditions Spatial Presence Questionnaire (MEC-SPQ) (Vorderer et al., 2004) – 32-64 item versions, 8 subscales</li> <li>• The Secondary Task Reaction Time (STRT) to assess non-mediated world attention</li> <li>• Eye-tracking can be used to investigate attentional attributes</li> </ul> <p>According to Bowman and McMahan (2007), the following elements have an impact on immersion from a technological standpoint regarding virtual reality applications:</p> <ul style="list-style-type: none"> <li>• Field of view (FOV) – the size of the visual field that can be viewed instantaneously</li> <li>• Field of regard (FOR) – total size of the visual field surrounding the user</li> <li>• Display size and display resolution</li> <li>• Stereoscopy – display of different images to each eye to provide an additional depth cue</li> <li>• Head-based rendering – the display of images based on the physical position and orientation of the user's head (produced by head tracking)</li> <li>• Realism of lighting</li> <li>• Framerate (e.g., FPS) and refresh rate (e.g., Hz)</li> </ul>
<b>Recommended Literature</b>	Agrawal, S., Simon, A., Bech, S., Bæntsen, K., & Forchhammer, S. (2020). Defining immersion: Literature review and implications for research on audiovisual experiences. <i>Journal of the Audio Engineering Society</i> , 68(6), 404–417. <a href="https://doi.org/10.17743/jaes.2020.0039">https://doi.org/10.17743/jaes.2020.0039</a>

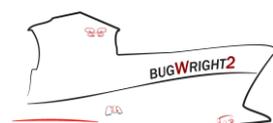


## References

- Agrawal, S., Simon, A., Bech, S., Bæntsen, K., & Forchhammer, S. (2020). Defining immersion: Literature review and implications for research on audiovisual experiences. *Journal of the Audio Engineering Society*, 68(6), 404–417. <http://www.aes.org/e-lib/browse.cfm?elib=20857>
- Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, 40(7), 36–43. <https://doi.org/10.1109/mc.2007.257>
- Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, 10(3), 282–297. <https://doi.org/10.1162/105474601300343612>
- Salanitri, D., Lawson, G., & Waterfield, B. (2016). The relationship between presence and trust in virtual reality. In *Proceedings of the European Conference on Cognitive Ergonomics* (pp. 1–4). ACM. <https://doi.org/10.1145/2970930.2970947>
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1535), 3549–3557. <https://doi.org/10.1098/rstb.2009.0138>
- Slater, M., Spanlang, B., & Corominas, D. (2010). Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. In T. DeRose & H. Hoppe (Eds.), *ACM SIGGRAPH 2010 papers on SIGGRAPH '10* (p. 1). ACM Press. <https://doi.org/10.1145/1833349.1778829>
- Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Laarni, J., Ravaja, N., Sacau, A., Baumgartner, T., & Jäncke, P. (2004). *MEC Spatial Presence Questionnaire (MEC-SPQ): Short Documentation and Instructions for Application*. Report to the European Community, Project Presence: MEC (IST-2001-37661). <https://academic.csuohio.edu/kneuendorf/frames/MECFull.pdf>

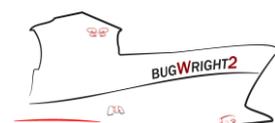


## Factsheet 9: Measuring Task Performance in Human-Robot Teams



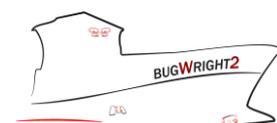
## Factsheet 9: Measuring Task Performance in Human-Robot Teams

<b>Keywords</b>	Task performance; human-computer interaction
<b>At a Glance</b>	There are several ways to measure task performance in work-related tasks. Popular options are the degree of correctness (e.g., error rate, false positives) or time (e.g., time on task, reaction time). While usually, either correctness or time is chosen for task performance measurement, a speed-accuracy trade-off must be considered. Performance measures need to fulfill the quality criteria of psychometric measurement (i.e., validity, reliability).
<b>Scientific Input</b>	
<b>Definition</b>	Human performance is the accomplishment of a task by a human operator or by a team of human operators (Gawron, 2019). Or, in more generalized terms, performance is the accomplishment of a task by a human operator or a team of two or more operators where team members can be human or robots (i.e., hybrid teams).
<b>Theoretical Background</b>	Performance measurement is crucial in deciding whether a human is good at conducting a task or not. However, due to the use of robotic agents in BUGWRIGHT2, it is necessary to not only focus on human performance but also on robot performance and the performance of the system (Kaupp & Makarenko, 2008). For example, automated robot path planning has been shown to increase system performance (Lewis et al., 2011). Human performance is closely related to mental load and other individual factors (Szalma & Teo, 2010).
<b>Recommendations for BUGWRIGHT2</b>	
<b>Measurement, Promotion, Intervention, Etc.</b>	<p>According to Gawron (2019), human performance metrics may be divided into six categories, of which the three most relevant to BUGWRIGHT2 are listed below:</p> <p><i>Accuracy</i> is the first category. Measures of accuracy include a correctness score, number of correct answers, percentage of correct answers, percentage of correct detections, and probability of correct detections section. Errors can also be used to measure accuracy. Error measures include absolute errors, average range scores, deviations, error rate, false alarm rate, number of errors, percentage of errors, and root mean square error.</p> <p><i>Time</i> is the second category of human performance indicators. Measures of time assume that tasks have a well-defined beginning and end so that the duration of task performance can be measured. Measures in this category include dichotic listening, detection time, glance duration, marking speed, movement time, reaction time, reading speed, search time, task load, and time to complete.</p> <p><i>Team performance</i> measurements are the third category. While Gawron (2019) only relates this to humans-only teams, in BUGWRIGHT2 this gains more relevance for hybrid teams consisting of humans and robots. Generally, team performance measurements can be understood as a combination of the aforementioned measures applied to team members, whereas Harriott and Adams (2013) try modeling human performance to estimate the overall system's performance when the robotic parameters are known.</p> <p>Possible performance metrics for human-robot teams are listed in more detail in Steinfeld et al. (2006). They are distinguished into task metrics (e.g., Novak et al., 2012), perception, management, manipulation, and robot performance. Most of the metrics for human performance can be categorized like this. However, the authors introduce important categories for measuring robot performance: Self-awareness (e.g., To which extent is the robot capable of identifying its own situation and potential obstacles or similar nearby and when to invoke human intervention?), human awareness (e.g., Is the robot able to understand human ways of interaction?, Can the robot system adapt to changes in human behavior if necessary?), and autonomy.</p>
<b>Recommended Literature</b>	Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., & Goodrich, M. (2006). Common metrics for human-robot interaction. In M. A. Goodrich (Ed.), <i>ACM Conferences, Proceedings of the 1<sup>st</sup> ACM SIGCHISIGART conference on Human-robot interaction</i> (p. 33). ACM. <a href="https://doi.org/10.1145/1121241.1121249">https://doi.org/10.1145/1121241.1121249</a>

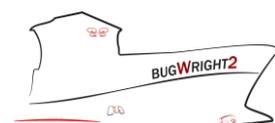


## References

- Gawron, V. J. (2019). *Human performance and situation awareness measures* (Third edition). CRC Press/Taylor & Francis Group.
- Harriott, C. E., & Adams, J. A. (2013). Modeling human performance for human-robot systems. *Reviews of Human Factors and Ergonomics*, 9(1), 94–130. <https://doi.org/10.1177/1557234X13501471>
- Kaupp, T., & Makarenko, A. (2008). Measuring human-robot team effectiveness to determine an appropriate autonomy level. In *2008 IEEE International Conference on Robotics and Automation* (pp. 2146–2151). IEEE. <https://doi.org/10.1109/ROBOT.2008.4543524>
- Lewis, M., Wang, H., Chien, S. Y., Velagapudi, P., Scerri, P., & Sycara, K. (2011). Process and performance in human-robot teams. *Journal of Cognitive Engineering and Decision Making*, 5(2), 186–208. <https://doi.org/10.1177/1555343411409323>
- Novak, D., Mihelj, M., & Munih, M. (2012). Dual-task performance in multimodal human-computer interaction: A psychophysiological perspective. *Multimedia Tools and Applications*, 56(3), 553–567. <https://doi.org/10.1007/s11042-010-0619-7>
- Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., & Goodrich, M. (2006). Common metrics for human-robot interaction. In M. A. Goodrich (Ed.), *ACM Conferences, Proceedings of the 1<sup>st</sup> ACM SIGCHISIGART conference on Human-robot interaction* (p. 33). ACM. <https://doi.org/10.1145/1121241.1121249>
- Szalma, J. L., & Teo, G. W. (2010). The Joint Effect of Task Characteristics and Neuroticism on the Performance, Workload, and Stress of Signal Detection. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 54(14), 1052–1056. <https://doi.org/10.1177/154193121005401401>

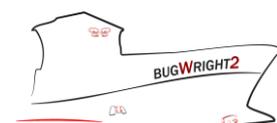


## Factsheet 10: Promoting Appropriate Trust in HRI through Design



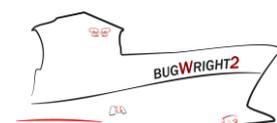
## Factsheet 10: Promoting Appropriate Trust in HRI through Design

<b>Keywords</b>	Trust promotion; design; human-robot interaction; multi-robot-systems
<b>At a Glance</b>	Interface design impacts trust in human-robot interaction (HRI). Trust in HRI is influenced by various variables such as high usability, adaptive levels of autonomy (LoA), polite communication, accurate and comprehensive feedback, or explanations of the robot's reliability and functionality. It is important to aim at calibrated trust levels and to avoid mistrust or overtrust.
<b>Scientific Input</b>	
<b>Short Introduction</b>	Robot factors are more important for trust in HRI than human factors (Hancock et al., 2011). For trustworthy HRI, based on Hoff and Bashir (2015), five design features are vital that are focussed on in the following sections. See Factsheet 14 for a general overview of the concept of trust in HRI.
<b>1/5 Appearance</b>	Increasing the humanness of robotic agents, for example, including a picture of an expert in the interface, might foster trust (de Visser et al., 2018; Pak et al., 2012). Also, keeping the age and gender of the robotic agent close to the human operator might help to build trust in HRI (Bass et al., 2013; Pak et al., 2014). However, anthropomorphic design can also have negative effects on trust as emotional relations can lead to a miscalibration of trust (Culley & Madhavan, 2013; Lee & See, 2004).
<b>2/5 Ease of Use</b>	Increased usability through higher visual clarity, among others (e.g., better traceability of results), goes along with increased trust (Atoyan et al., 2006). Wang et al. (2011) found trust-promoting effects of higher salience of automated feedback (Hoff & Bashir, 2015).
<b>3/5 Communication Style</b>	A polite way of communication that is non-interruptive and patient is trustworthy (Parasuraman & Miller, 2004; Spain & Madhavan, 2009). Regarding the type of communication, research is heterogeneous. Lee and See (2002) argue for the advantages of speech-based interfaces. Here, subtle auditory cues like inflection, sighs, and pauses can signal uncertainty, which makes use of people's ability to understand emotional hints. Other researchers prefer text-based communication (Gong, 2008; Lee, 2008).
<b>4/5 Transparency/ Feedback</b>	<i>Accuracy.</i> Accurate feedback on the robotic agent's reliability and functionality (e.g., explanations for failures) fosters appropriately high trust (Dadashi et al., 2013; Dzindolet et al., 2003; Gao & Lee, 2006; Wang et al., 2009) and task performance (Bagheri & Jamieson, 2004; Bean et al., 2011). <i>Timing.</i> Explanations of robot behaviors ahead of time ("feedforward") are crucial for keeping steady trust in case of automation failures (Desai et al., 2013; Haspiel et al., 2018). <i>Form of explanation:</i> Comprehensive visualizations that include symbolic and haptic representations of the robot's internal decisions were experienced as trustworthy (Edmonds et al., 2019). Semantic symbols like smileys led to more drastic trust changes (in both directions) than non-semantic symbols (Desai et al., 2013). Small intentional failures of the automation ("adaptive failure") could recalibrate the human's trust to a lower, more appropriate level in case of overreliance (Chen et al., 2020). "Trust Repair Actions" by the robot, like offering apologies, suggesting different solutions, giving explanations, and privacy guarantees might help in case of lack of reliance (de Visser et al., 2018).
<b>5/5 Level of Control</b>	Trust in HRI is higher with <i>mixed LoA</i> compared to fully autonomous systems and swarms (Huaio, 2019; Verberne et al., 2012). However, time delays in "mixed" situations (e.g., the system is waiting for the human to react, Verberne et al., 2012) can be detrimental in time-critical situations. <i>Adaptive automation</i> switches between different LoAs to fit different situations and different users with their own preferences or characteristics (Thropp, 2006).
<b>Recommendations for BUGWRIGHT2</b>	
<b>Measurement, Promotion, Intervention, Etc.</b>	For a trustworthy robotic system we recommend to a) focus on the mechanical robot appearance, though anthropomorphic appearance might not be relevant in the marine sector, b) ensure high usability (i.e., high visual clarity, high salience of feedback), c) use polite text-based communication (i.e., non-interruptive, patient), d) include comprehensive and graphic (non-semantic) explanations of robots' functionality and reliability, as well as warnings ahead of time when (and why) failures seem likely, and e) choose mixed or adaptive LoA appropriate for the specific situation.
<b>Recommended Literature</b>	Desai, M., Kaniarasu, P., Medvedev, M., Steinfeld, A., & Yanco, H. (2013). Impact of robot failures and feedback on real-time trust. In <i>HRI 2013 Proceedings</i> , (pp. 251-258). <a href="https://doi.org/10.1109/HRI.2013.6483596">https://doi.org/10.1109/HRI.2013.6483596</a> Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. <i>Human Factors</i> , 57(3), 407-434. <a href="https://doi.org/10.1177/0018720814547570">https://doi.org/10.1177/0018720814547570</a>

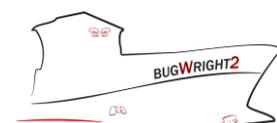


## References

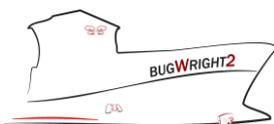
- Atoyan, H., Duquet, J. R., & Robert, J. M. (2006). Trust in new decision aid systems. *Proceedings of the 18<sup>th</sup> International Conference of the Association Francophone d'Interaction Homme-Machine* (pp. 115–122). ACM. <https://doi.org/10.1145/1132736.1132751>
- Bagheri, N., & Jamieson, G. A. (2004). The impact of context-related reliability on automation failure detection and scanning behaviour. *2004 IEEE International Conference on Systems, Man and Cybernetics, 1*, 212–217.
- Bass, B., Goodwin, M., Brennan, K., Pak, R., & McLaughlin, A. (2013). Effects of age and gender stereotypes on trust in an anthropomorphic decision aid. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*(1), 1575-1579. <https://doi.org/10.1177/1541931213571351>
- Bean, N. H., Rice, S. C., & Keller, M. D. (2011). The effect of gestalt psychology on the system-wide trust strategy in automation. In *Proceedings of the Human Factors and Ergonomics Society 55<sup>th</sup> Annual Meeting* (pp. 1417–1421). Human Factors and Ergonomics Society. <https://doi.org/10.1177/1071181311551295>
- Chen, M., Nikolaidis, S., Soh, H., Hsu, D., & Srinivasa, S. (2020). Trust-aware decision making for human-robot collaboration: Model learning and planning. *ACM Transactions on Human-Robot Interaction, 9*(2), Article 9. <https://doi.org/10.1145/3359616>
- Culley, K. E., & Madhavan, P. (2013). A note of caution regarding anthropomorphism in HCI agents. *Computers in Human Behavior, 29*(3), 577-579. <https://doi.org/10.1016/j.chb.2012.11.023>
- Dadashi, N., Stedmon, A. W., & Pridmore, T. P. (2013). Semi-automated CCTV surveillance: The effects of system confidence, system accuracy and task complexity on operator vigilance, reliance and workload. *Applied Ergonomics, 44*(5), 730-738. <https://doi.org/10.1016/j.apergo.2012.04.012>
- Desai, M., Kaniarasu, P., Medvedev, M., Steinfeld, A., & Yanco, H. (2013). Impact of robot failures and feedback on real-time trust. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 251–258). IEEE. <https://doi.org/10.1109/HRI.2013.6483596>
- de Visser, E. J., Pak, R., & Shaw, T. H. (2018). From 'automation' to 'autonomy': The importance of trust repair in human-machine interaction. *Ergonomics, 61*(10), 1409-1427. <https://doi.org/10.1080/00140139.2018.1457725>
- Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies, 58*(6), 697-718. [https://doi.org/10.1016/S1071-5819\(03\)00038-7](https://doi.org/10.1016/S1071-5819(03)00038-7)
- Edmonds, M., Gao, F., Liu, Hangxin, Xie, X., Siyuan, Q., Rothrock, B., Zhu, Y., Wu, Y. N., Lu, H., & Zhu, S. C. (2019). A tale of two explanations: Enhancing human trust by explaining robot behavior. *Science Robotics, 4*(37), Article eaay4663. <https://doi.org/10.1126/scirobotics.aay4663>
- Gao, J., & Lee, J. D. (2006). Effect of Shared information on trust and reliance in a demand forecasting task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(3), 215–219. <https://doi.org/10.1177/154193120605000302>
- Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior, 24*(4), 1494–1509. <https://doi.org/10.1016/j.chb.2007.05.007>
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*(5), 517-527. <https://doi.org/10.1177/0018720811417254>



- Haspiel, J., Du, N., Meyerson, J., Robert, L. P., Jr., Tilbury, D., Yang, X. J., & Pradhan, A. K. (2018). Explanations and expectations: Trust building in automated vehicles. In *HRI Companion, March 5-8, 2018, Chicago, IL USA*, 119-120. <https://doi.org/10.1145/3173386.3177057>
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407-434. <https://doi.org/10.1177/0018720814547570>
- Huao, L. (2019). *A computational model of human trust in supervisory control of robotic swarms* [Unpublished Master's Thesis]. University of Pittsburgh.
- Lee, E.-J. (2008). Flattery may get computers somewhere, sometimes: The moderating role of output modality, computer gender, and user gender. *International Journal of Human-Computer Studies*, 66(11), 789-800. <https://doi.org/10.1016/j.ijhcs.2008.07.009>
- Lee, J. D., & See, K. A. (2002). Trust in computer technology and the implications for design and evaluation. *AAAI Technical Report FS-02-02*. AAAI.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50-80. [https://doi.org/10.1518/hfes.46.1.50\\_30392](https://doi.org/10.1518/hfes.46.1.50_30392)
- Pak, R., McLaughlin, A. C., & Bass, B. (2014) A multi-level analysis of the effects of age and gender stereotypes on trust in anthropomorphic technology by younger and older adults. *Ergonomics*, 57(9), 1277-1289. <https://doi.org/10.1080/00140139.2014.928750>
- Pak, R., Fink, N., Price, M., Bass, B., & Sturre, L. (2012). Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics*, 55(9), 1059-1072. <https://doi.org/10.1080/00140139.2012.691554>
- Parasuraman, R., & Miller, C. A. (2004). Trust and etiquette in high-criticality automated systems. *Communications of the ACM*, 47(4), 51-55. <https://doi.org/10.1145/975817.975844>
- Spain, R. D., & Madhavan, P. (2009). The role of automation etiquette and pedigree in trust and dependence. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53(4), 339-343. <https://doi.org/10.1177/154193120905300437>
- Thropp, J. E. (2006). *Individual preferences using automation* (Unpublished doctoral dissertation). University of Central Florida, Orlando.
- Verberne, F. M., Ham, J., & Midden, C. J. (2012). Trust in smart systems sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars. *Human Factors*, 54(5), 799-810. <https://doi.org/10.1177/0018720812443825>
- Wang, L., Jamieson, G. A., & Hollands, J. G. (2009). Trust and reliance on an automated combat identification system: The role of aid reliability and reliability disclosure. *Human Factors*, 51(3), 281-291. <https://doi.org/10.1177/0018720809338842>
- Wang, L., Jamieson, G. A., & Hollands, J. G. (2011). The effects of design features on users' trust in and reliance on a combat identification system. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 55(1), 375-379. <https://doi.org/10.1177/1071181311551077>

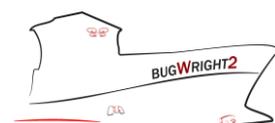


## **Factsheet 11: Self-Efficacy in the Context of Human-Robot Interaction**



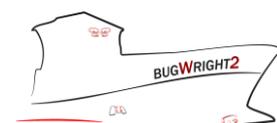
## Factsheet 11: Self-Efficacy in the Context of Human-Robot Interaction

<b>Keywords</b>	Self-efficacy; human-robot interaction; competence belief; technology acceptance; human factor; performance expectancy; effort expectancy
<b>At a Glance</b>	Self-efficacy (SE) describes the perceived confidence to successfully master a specific situation or task. SE related to robots and the use of technology influences its use motivation, evaluation, and perseverance when interacting with them. To foster SE, four central factors act as starting points for SE interventions including mastery experiences.
<b>Scientific Input</b>	
<b>Definition</b>	SE describes one's self-perceived confidence to succeed in a situation or task. SE is domain-specific, meaning that people do not have one overarching SE, but rather SE is a complex set of self-beliefs related to SE in different areas of functioning (Marsh et al., 2017) such as the use of information and communication technology (ICT, e.g., Brosnan, 1998) and robotics (e.g., Rosenthal-von der Pütten & Bock, 2018). SE does not equal efficacy, as people might possess all necessary abilities to succeed in a situation but do not believe that they might be able to do so. Thus, SE is a substantial predictor of use motivation and behavior (Bandura, 2006).
<b>Technology Acceptance and SE</b>	SE is integrated into widely used models of the acceptance and use of technology (for more information, see Venkatesh et al., 2016). Here, the research focus was on SE related to computers. Recent research provides evidence for the role of SE in robot acceptance and human-robot interaction (HRI), as well (Bröhl et al., 2019). SE predicts performance expectancy (i.e., perceived usefulness) and effort expectancy (i.e., perceived ease of use). Performance and effort expectancy, in turn, are two key determinants of future use behavior or recommendation behavior. In general, high SE results in more positive robot evaluation (Rosenthal-von der Pütten & Bock, 2018). However, existing research is promising but limited as SE in robotics is a new research field, mostly focused on single service robotics (e.g., in the health sector) not on robotics in industry and production, or robotic swarms (Savela et al., 2018).
<b>Development of SE</b>	SE is learned and, thus, changes by experiences people make. Four central factors influence SE (Bandura, 1977, 2006). Most importantly, <i>mastery experience</i> ; the more you gain experience in a specific situation and manage to be successful, the more likely you will expect to be successful again in a similar challenge. If direct experience is not possible, it is also feasible to promote SE by experiencing someone else's mastery of the situation, called <i>vicarious experience</i> . Reading a factsheet about robots leads to higher SE in elderly people than interaction with the robot itself (Zafari et al., 2019). The third factor is <i>social persuasion</i> . Direct encouragement or discouragement from others influences SE. Thereby, social persuasion can come from other people but also from the robotic agent itself. Zafari et al. (2019) showed that in the context of social interaction, interacting with a robot that gives person-oriented feedback (praising, encouraging) was associated with higher SE and a less frustrating experience while interacting with a task-oriented robot was not. Lastly, experiencing <i>physiological symptoms</i> like heart racing or sweating before a challenging situation can further decrease SE in people that believe these symptoms to be signs of their low ability.
<b>Recommendations for BUGWRIGHT2</b>	
<b>Measurement and Promotion</b>	To reach a high robot acceptance, high SE of employees (e.g., surveyor) directly in contact with the remote inspection technologies (RITs) is crucial. When <i>tracking</i> SE, different SE domains are to be considered, such as operating the robot and monitoring the inspection. <i>Promoting</i> SE, a consideration of all four sources of SE identified by Bandura (1977) is necessary. Tutorials and training sessions (mastery experiences), praise and encouragement, as well as instructions and explanations (by an instructor or the robot itself), and demonstration or modeling (vicarious experience) help to guarantee high SE related to the new robotic system.
<b>Recommended Literature</b>	Rosenthal-von der Pütten, A. M., & Bock, N. (2018). Development and validation of the self-efficacy in human-robot-interaction scale (SE-HRI). <i>ACM Transactions on Human-Robot Interaction</i> , 7(3), 1–30. <a href="https://doi.org/10.1145/3139352">https://doi.org/10.1145/3139352</a>

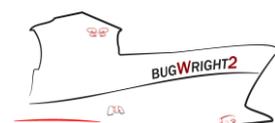


## References

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Information Age Publishing.
- Bröhl, C., Nelles, J., Brandl, C., Mertens, A., & Nitsch, V. (2019). Human–robot collaboration acceptance model: Development and comparison for Germany, Japan, China and the USA. *International Journal of Social Robotics*, 11(5), 709–726. <https://doi.org/10.1007/s12369-019-00593-0>
- Brosnan, M. J. (1998). The impact of computer anxiety and self-efficacy upon performance. *Journal of Computer Assisted Learning*, 14(3), 223–234. <https://doi.org/10.1046/j.1365-2729.1998.143059.x>
- Marsh, H. W., Martin, A. J., Yeung, A. S., & Craven, R. G. (2017). Competence self-perceptions. In A. J. Elliot (Ed.), *Handbook of competence and motivation* (pp. 85–115). Guilford Press.
- Rosenthal-von der Pütten, A. M., & Bock, N. (2018). Development and validation of the self-efficacy in human-robot-interaction scale (SE-HRI). *ACM Transactions on Human-Robot Interaction*, 7(3), 1–30. <https://doi.org/10.1145/3139352>
- Savela, N., Turja, T., & Oksanen, A. (2018). Social acceptance of robots in different occupational fields: A systematic literature review. *International Journal of Social Robotics*, 10(4), 493–502. <https://doi.org/10.1007/s12369-017-0452-5>
- Venkatesh, V., Thong, J., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, 17(5), 328–376. <https://doi.org/10.17705/1jais.00428>
- Zafari, S., Schwaninger, I., Hirschmanner, M., Schmidbauer, C., Weiss, A., & Koeszegi, S. T. (2019). "You are doing so great!" – The effect of a robot's interaction style on self-efficacy in HRI. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (pp. 1–7). IEEE. <https://doi.org/10.1109/RO-MAN46459.2019.8956437>

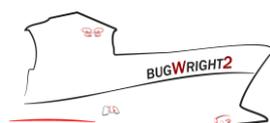


## Factsheet 12: Situation Awareness – Knowing what's going on



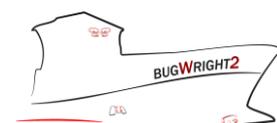
## Factsheet 12: Situation awareness – Knowing what’s going on

<b>Keywords</b>	Situation awareness; decision making
<b>At a Glance</b>	Situation Awareness (SA) reflects the knowledge about the current circumstances a person is in. More precisely, it reflects the portion from this knowledge that is relevant for conducting the task that the individual is performing. SA has been identified as a crucial but still elusive foundation for successful decision-making in a variety of fields. It has been determined that lack of SA is one of the key causes of human error-caused incidents (Endsley et al., 2003; Nullmeyer et al. 2005).
<b>Scientific Input</b>	
<b>Definition</b>	SA has been studied from three different angles: SA states, SA systems, and SA processes. SA states refer to a person’s genuine understanding of the situation. It is the distribution of SA in teams and between objects within the environment and the exchange of SA between system parts that are referred to as SA systems in this context. SA processes refer to the updating of SA states, as well as to what leads to the moment-to-moment change of SA in real-time (Endsley et al., 2003; Lundberg, 2015).
<b>Theoretical Background</b>	SA first occurred in the technical literature in 1983 while explaining the benefits of a prototype touch-screen navigation display (Biferno & Stanley, 1983). Integrated “vertical-situation” and “horizontal-situation” displays were created for commercial aircrafts in the early 1980s to replace numerous electro-mechanical instruments. The situation displays integrated information from several instruments allowing for more efficient access to important flying data, enhancing SA, and decreasing pilot workload. Individuals and teams must be able to perform well in their environment; therefore, SA has a wide range of applications. As a result, SA has been used in different fields of work and is being included in several different research settings.
<b>Recommendations for BUGWRIGHT2</b>	
<b>Variety of Measurements</b>	<p>As SA is especially relevant for coordination at the hull during the inspection and for interacting with any robotic system appropriately, it is relevant to BUGWRIGHT2 to control for the SA of workers involved in the inspection. Three method categories for measurement are described below.</p> <p><i>Objective measures.</i> By comparing an individual’s perceptions of a situation or environment to some “ground truth” reality, objective measurements directly assess SA. Objective measures in particular take data from the subject about his or her perception of the situation and compare it to what is happening to assess the accuracy of their SA at any given time. As a result, this sort of assessment offers a direct measure of SA and eliminates the need for operators or observers to make situational knowledge judgments based on limited data. Objective measures can be collected in one of three ways: in real-time as the task is done (Jones &amp; Endsley, 2000), during a task interruption (Endsley, 1995), or after the work has been finished with a post-test.</p> <p><i>Subjective Measures.</i> Ask people to rate their own or others’ SA on a scale, e.g., the participant situation awareness questionnaire (PSAQ, Strater et al., 2001) or the situation awareness rating method (SART, Taylor, 2017). Experienced observers can make subjective estimates of an individual’s SA (e.g., peers, commanders, or trained external experts). Because the observer has more information about the real state of the environment than the operator, these observer evaluations may be superior to SA self-ratings (i.e., trained observers may have more complete knowledge of the situation).</p> <p><i>Performance and behavioral measures.</i> As better performance suggests better SA, performance metrics “infer” SA from the result (i.e., task performance outcomes). Quantity of output or productivity level, time to complete the job or respond to an event, and correctness of the answer or, conversely, the number of errors made are all common performance indicators. The major benefit of performance measurements is that they can be gathered objectively and without interfering with job execution.</p>
<b>Recommended Literature</b>	Gawron, V. J. (2019). <i>Human performance and situation awareness measures</i> . CRC Press. <a href="https://doi.org/10.1201/9780429001024">https://doi.org/10.1201/9780429001024</a>



## References

- Biferno, M. A., & Stanley, D. L. (1983). The touch-sensitive control/display unit: A promising computer interface. In *SAE Technical Paper Series, SAE Technical Paper Series*. SAE International 400 Commonwealth Drive, Warrendale, PA, United States. <https://doi.org/10.4271/831532>
- Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 65–84. <https://doi.org/10.1518/001872095779049499>
- Endsley, M. R., Bolte, B., & Jones, D. G. (2003). *Designing for Situation Awareness*. CRC Press. <https://doi.org/10.1201/9780203485088>
- Gawron, V. J. (2019). *Human performance and situation awareness measures*. CRC Press. <https://doi.org/10.1201/9780429001024>
- Jones, D. G., & Endsley, M. R. (2000). Examining the validity of real-time probes as a metric of situation awareness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 44(1), 278–279. <https://doi.org/10.1177/154193120004400175>
- Lundberg, J. (2015). Situation awareness systems, states and processes: A holistic framework. *Theoretical Issues in Ergonomics Science*, 16(5), 447–473. <https://doi.org/10.1080/1463922X.2015.1008601>
- Nullmeyer, R., Stella, L. C. D., & Montijo, G. (2005). Mishap trends in several air force aircraft: Implications for CRM training. *2005 International Symposium on Aviation Psychology*, 534–539. <https://corescholar.libraries.wright.edu/isap2005/82>
- Strater, L. D., Endsley, M. R., Pleban, R. J., & Matthews, M. D. (2001). *Measures of Platoon Leader Situation Awareness in Virtual Decision-Making Exercises*. Defense Technical Information Center. <https://doi.org/10.21236/ada393977>
- Taylor, R. M. (2017). Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design. In E. Salas (Ed.), *Situational Awareness* (1st ed., pp.111–128). Routledge. <https://doi.org/10.4324/9781315087924-8>

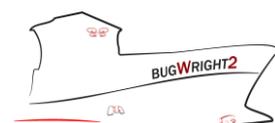


## Factsheet 13: SMART Work via SMART Work Design



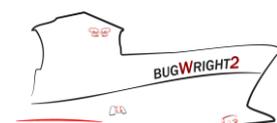
## Factsheet 13: SMART Work via SMART Work Design

<b>Keywords</b>	SMART work; work design; motivational work; work characteristics
<b>At a Glance</b>	SMART work refers to a concept for designing meaningful and motivating work, characterized by five key work characteristics (i.e., stimulation, mastery, agency, relations, tolerable demands) that positively influence individuals' commitment, creativity, engagement, performance, and innovation.
<b>Scientific Input</b>	
<b>Definition and Theoretical Background</b>	Work design refers to "the content and organization of one's work tasks, activities, relationships, and responsibilities" (Parker, 2014, p. 662). From a psychological perspective, work should be designed to enable learning and development, facilitate well-being and health, and cope with ambidexterity (e.g., simultaneously achieving control and flexibility, Parker, 2014). Decades of research on work design identified characteristics that impact employees' motivation, well-being, and performance (e.g., Hackman & Oldham, 1976; Parker et al., 2017). The SMART approach unifies this literature into a simple model, focusing on five core work features. The acronym SMART stands for work that involves stimulation, mastery, agency, relations, and tolerable demands (see Klonek & Parker, 2021; Parker et al., 2017).
<b>Stimulating Mastery Agency Relational Tolerable demands</b>  (see Klonek & Parker, 2021; Parker, 2014; Parker et al., 2017)	<p><i>Stimulating.</i> SMART jobs enable the use of an adequate <i>skill variety</i> in a wide range of tasks (<i>task variety</i>) and request to think outside the box (<i>problem-solving demands</i>). Stimulating work leads to more creativity and innovation.</p> <p><i>Mastery.</i> SMART jobs provide individuals with realistic information about the given task and roles (i.e., role clarity). They include (performance-related) <i>feedback</i>. <i>Task identity</i> is high, allowing task completion from the beginning to the end. Mastery experiences cultivate self-efficacy (see Factsheet 11).</p> <p><i>Agency.</i> SMART jobs are characterized by high autonomy and the decision-making of the individuals involved. This includes, among others, control over scheduling (e.g., when to accomplish the task) and work methods (e.g., how to accomplish the task). Low agency corresponds with high risks for mental health issues.</p> <p><i>Relational.</i> SMART jobs feature <i>social support</i> (i.e., from colleagues or supervisors). The work tasks are purposeful in relation to others and society (<i>task significance</i>). Individuals perceive to be valued for what they do (<i>social worth</i>). Relational work boosts commitment to the organization.</p> <p><i>Tolerable demands.</i> SMART jobs request moderate and manageable levels of <i>time pressure</i> and <i>workload</i> but also regarding <i>emotional demands</i>. Inconsistency of feedback and conflicting task-related information are low (<i>low role conflict</i>). Tolerable demands are crucial to reducing involuntary absence through (mental or physical) illness (Schaufeli et al., 2009).</p>
<b>Recommendations for BUGWRIGHT2</b>	
<b>Evaluating Work Features</b>	The implementation of automated robotics in hull inspections will lead to massive changes in the work processes, qualitative changes in team interactions, and the transformation of job profiles and roles. The SMART work design approach offers a helpful and simple framework to evaluate work features that have a well-established impact on employees' motivation, well-being, and performance. The five SMART dimensions and their subdimensions may act as a guideline for a structured discussion on the future automated work environment because subjective perceptions are valid measures of job characteristics (see Parker, 2014) in addition to international criteria of functional or dysfunctional work design (see DIN EN ISO 9241-2).
<b>Recommended Literature</b>	<p>Parker, S. K. (2014). Beyond motivation: Job and work design for development, health, ambidexterity, and more. <i>Annual Review of Psychology</i>, 65, 661-691. <a href="https://doi.org/10.1146/annurev-psych-010213-115208">https://doi.org/10.1146/annurev-psych-010213-115208</a></p> <p>Parker, S. K., Morgeson, F. P., &amp; Johns, G. (2017). One hundred years of work design research: Looking back and looking forward. <i>Journal of applied psychology</i>, 102(3), 403. <a href="https://doi.org/10.1037/apl000106">https://doi.org/10.1037/apl000106</a></p>



## References

- Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, 16(2), 250–279. [https://doi.org/10.1016/0030-5073\(76\)90016-7](https://doi.org/10.1016/0030-5073(76)90016-7)
- Klonek, F., & Parker, S. K. (2021). Designing smart teamwork: How work design can boost performance in virtual teams. *Organizational Dynamics*, 50(1), 100841. <https://doi.org/10.1016/j.orgdyn.2021.100841>
- Parker, S. K. (2014). Beyond motivation: Job and work design for development, health, ambidexterity, and more. *Annual Review of Psychology*, 65(1), 661–691. <https://doi.org/10.1146/annurev-psych-010213-115208>
- Parker, S. K., Morgeson, F. P., & Johns, G. (2017). One hundred years of work design research: Looking back and looking forward. *The Journal of Applied Psychology*, 102(3), 403–420. <https://doi.org/10.1037/apl0000106>
- Schaufeli, W. B., Bakker, A. B., & van Rhenen, W. (2009). How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *Journal of Organizational Behavior*, 30(7), 893–917. <https://doi.org/10.1002/job.595>

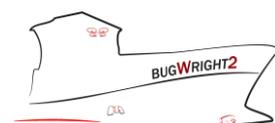


## Factsheet 14: Trust in Automation in a Multi-Robot System



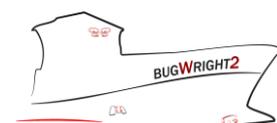
## Factsheet 14: Trust in Automation in a Multi-Robot System

<b>Keywords</b>	Trust; human-robot interaction; multi-robot systems
<b>At a Glance</b>	Trust in automation is an important key factor to ensure technology acceptance and good human-robot team (HRT) performance. Trust in automation is influenced by the reliability, explainability, as well as design of robots, which results in important implications for robot and system design.
<b>Scientific Input</b>	
<b>Definition and Theoretical Background</b>	Trust in automation is defined as “the attitude that an agent will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability” (Lee & See, 2004, p. 51). This describes, on the one hand, a person’s tendency to trust robots, and on the other hand the willingness to accept some vulnerability and risk by trusting them with certain actions (Lee & See, 2004). This trust depends on the performance, the understanding of the inner workings, and the purpose of the automated system (Lee & Moray, 1992). Hoff and Bashier (2015) introduced a trust model which includes three layers of trust that are each influenced by different factors: dispositional trust (i.e., the personal enduring tendency to trust), context-dependent situational trust, and learned trust, which is based on past experiences with a specific system.
<b>Empirical Relations</b>	Trust in automation is important for technology acceptance and use (Meeßen et al., 2020), especially in circumstances where the adaption of the technology is not voluntary (Schauffel & Ellwart, 2021). Trust also increases HRT performance (You & Robert, 2019) and is a precondition of beneficial effects of using automated systems like higher well-being as well as better decision quality (Hertel et al., 2019). It is important to note though that too much trust in automation can be as bad as too little trust in automation. Instead, we should aim for a realistic, well-calibrated amount of trust appropriate to the robot’s capabilities to avoid misuse or disuse of the automation (Ososky et al., 2013). What influences trust in automation? A meta-analysis found that the operator’s emotive states (e.g., attitudes) are highly relevant for trust (Schaefer et al., 2016). Another meta-analysis by Hancock et al. (2011), however, concludes that other factors are more important for HRT. Here, environmental factors (e.g., team collaboration) but above all the robot’s performance and attributes like type, size, or behavior have the biggest influence on trust in automation. Robots’ reliability is most crucial (Hancock et al., 2011), as especially robot failures in early phases of human-robot interaction lower trust (Desai et al., 2013). When using multi-robot systems and robot swarms, the level of autonomy (LoA) has a big impact on trust dynamics since LoA influences how easily humans can evaluate the robot’s reliability (Lee & Moray, 1992). For example, only a high LoA reliability can foster trust because humans are only able to actively “evaluate” robots’ performance in this condition. Further, in supervisory tasks the physical swarm characteristics (i.e., coherence and concentration) tend to be even more important for human-swarm trust than robot performance (i.e., number of targets found, Huao, 2019).
<b>Recommendations for BUGWRIGHT2</b>	
<b>Measurement, Promotion, Intervention, Etc.</b>	Trust in automation is measured in several ways in existing research. For BUGWRIGHT2, we recommend using questionnaires that rate the predictability, reliability, and faith in the robot (Muir, 1990), or to measure the dispositional trust of operators (Edmonds et al., 2019). Another trust indicator to measure situational trust is to examine when the human operator takes back the control from the automated system as an indicator for the lack of trust (Chen et al., 2020). To promote appropriate trust (i.e., calibrated and balanced: “not too low nor too high”), one could design the robot to monitor human trust and include intentional failures in case of human overreliance (Chen et al., 2020). Further, specific explanations of robot behavior <i>before</i> the automated system starts acting (Haspiel, 2018) promote trust (see Edmonds et al., 2019 for details). For further design recommendations (e.g., feedback, interface), see Factsheet 10 and Factsheet 13.
<b>Recommended Literature</b>	Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. <i>Human Factors</i> , 46(1), 50-80. <a href="https://doi.org/10.1518/hfes.46.1.50.30392">https://doi.org/10.1518/hfes.46.1.50.30392</a>

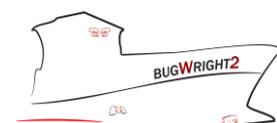


## References

- Chen, M., Nikolaidis, S., Soh, H., Hsu, D., & Srinivasa, S. (2020). Trust-aware decision making for human-robot collaboration: Model learning and planning. *ACM Transactions on Human-Robot Interaction*, 9(2), Article 9. <https://doi.org/10.1145/3359616>
- Desai, M., Kaniarasu, P., Medvedev, M., Steinfeld, A., & Yanco, H. (2013). Impact of robot failures and feedback on real-time trust. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 251–258). IEEE. <https://doi.org/10.1109/HRI.2013.6483596>
- Edmonds, M., Gao, F., Liu, Hangxin, Xie, X., Siyuan, Q., Rothrock, B., Zhu, Y., Wu, Y. N., Lu, H., & Zhu, S. C. (2019). A tale of two explanations: Enhancing human trust by explaining robot behavior. *Science Robotics* 4(37), Article eaay4663. <https://doi.org/10.1126/scirobotics.aay4663>
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, 53(5), 517–527. <https://doi.org/10.1177/0018720811417254>
- Haspiel, J., Du, N., Meyerson, J., Robert Jr., L. P., Tilbury, D., Yang, X. J., & Pradhan, A. K. (2018). Explanations and expectations: Trust building in automated vehicles. In T. Kanda, S. Šabanović, G. Hoffman, & A. Tapus (Eds.), *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI '18* (pp. 119–120). ACM Press. <https://doi.org/10.1145/3173386.3177057>
- Hertel, G., Meessen, S. M., Riehle, D. M., Thielsch, M. T., Nohe, C., & Becker, J. (2019). Directed forgetting in organisations: The positive effects of decision support systems on mental resources and well-being. *Ergonomics*, 62(5), 597–611. <https://doi.org/10.1080/00140139.2019.1574361>
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434. <https://doi.org/10.1177/0018720814547570>
- Hua0, L. (2019). *A computational model of human trust in supervisory control of robotic swarms* [Unpublished Master's Thesis]. University of Pittsburgh. <http://d-scholarship.pitt.edu/37400/>
- Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35(10), 1243–1270. <https://doi.org/10.1080/00140139208967392>
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80. <https://doi.org/10.1518/hfes.46.1.50.30392>
- Meeßen, S. M., Thielsch, M. T., & Hertel, G. (2020). Trust in Management Information Systems (MIS). *Zeitschrift für Arbeits- und Organisationspsychologie*, 64(1), 6–16. <https://doi.org/10.1026/0932-4089/a000306>
- Muir, B. M. (1990). *Operator's trust in and use of automatic controllers in a supervisory process control task*. University of Toronto.
- Osofsky, S., Schuster, D., Philipps, E., & Jentsch, F. (2013). Building appropriate trust in human-robot teams. In *Trust and Autonomous Systems: Papers from the 2013 AAAI Spring Symposium*, 60–65.
- 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: Implications for understanding autonomy in future systems. *Human Factors*, 58(3), 377–400. <https://doi.org/10.1177/0018720816634228>

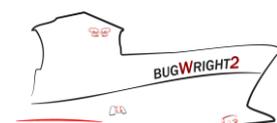


- Schauffel, N., & Ellwart, T. (2021). Forced virtuality during COVID-19: A multigroup perspective on technology acceptance of public digital services. *Zeitschrift für Arbeits- und Organisationspsychologie*, 65(4), 244-257. <https://doi.org/10.1026/0932-4089/a000366>
- You, S., & Robert, L. (2019). Trusting robots in teams: Examining the impacts of trusting robots on team performance and satisfaction. In T. Bui (Ed.), *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS 2019)* <https://doi.org/10.24251/HICSS.2019.031>



## Conclusion

The implementation of adaptable autonomous robotic solutions changes the world of work from all-human teams to human-robot teams. This change can be perceived as an opportunity for the humans involved, for example, when robotic solutions increase human safety and reduce human workload. However, robots at work can also be a threat if roles and responsibilities are ambiguous, new safety risks occur, or trust in the robotic system is miscalibrated (e.g., overtrust). In this e-book, we spotlighted central psychological factors for the success of human-robot teams from a scientific and practical perspective. The factsheets provide insights into psychological concepts about human-robot teams. Concrete recommendations provide starting points to design robotic solutions to be an opportunity, not a threat for humans involved. We hope this e-book is helpful not only for the BUGWRIGHT2 project but also for any interested reader dealing with the implementation of robotic solutions in interdisciplinary work contexts.



## The Authors



**Nathalie Schauffel** is a doctoral candidate in the unit Business Psychology at Trier University. Her research interests include human-agent teaming and humane socio-digital work design. She is also interested in academic and vocational self-concepts and the role of self-perceptions on the evaluation and acceptance of workplace technology.



**Jan Gründling** is a doctoral candidate in the unit Human-Computer Interaction at Trier University. His research interest is to find suitable concepts for the evaluation of VR environments of different, especially abstract types, and formulate resulting design guidelines for these environments. These can then be used for data analyses (immersive analytics) as in BUGWRIGHT2.



**Benjamin Ewerz** is a doctoral candidate in the unit of Business Psychology at Trier University. His research interests include the impact of user knowledge on human-agent teams. A focus here lies on areas of user knowledge that impact the development of trust and acceptance.



**Benjamin Weyers** is an assistant professor at Trier University and Head of the Human-Computer Interaction Group. He is interested in the research and development of interactive analysis methods for abstract and scientific data using immersive systems as well as the integration of VR/AR into the control of technical systems to support human users in semi-automated control scenarios.



**Thomas Ellwart** is a full professor and chair of the unit Business Psychology, which represents the field of work, organizational, and personnel psychology at Trier University. His research interests include socio-digital system design and human-agent teaming, focusing primarily on team processes and team cognition.

