Clap Twice for Trust: How Framing Shapes Perceived Trustworthiness in VR-mediated Industrial Human-Robot Interaction

Matouš Jelínek
matous@sdu.dk
University of Southern Denmark
Kolding, Denmark

Pat Treusch
pat@sdu.dk
University of Southern Denmark
Kolding, Denmark

Abstract

Trust is a crucial element in industrial human-robot interaction, yet little is known about how initial system descriptions shape perceived trustworthiness of the system, not only before use. Transparency about a system's capabilities is crucial for reaching appropriate reliance and perceived trustworthiness. This paper underlines the importance of what users are told about a system before their first contact. This information shapes their expectations about the robot and early reliance. Thus, this study investigates how the verbal framing of an industrial robot—either as a low-capability prototype or a high-capability advanced technology—shapes perceived trustworthiness of the system. Utilizing a VR simulation of a poultry packing robot, participants (N=21) interacted with an identical system. Notably, participants working with the high-capability framing rated the system as significantly more competent, showed higher reliability ratings, and also tended to express trusting behavior sooner than those interacting with the system framed as a low-capability prototype. The results show that communicated competences affect perceived trustworthiness and trusting behavior in a simulated, VR-mediated industrial human-robot interaction. Based on those findings, we argue for the relevance of proper framing, as, e.g., improper framing poses a risk of a miscalibration when system performance does not match its framing. Framing, thus, we argue, is an under-valued dimension of evaluating trustworthiness in (industrial) HRI.

CCS Concepts

• **Human-centered computing → Empirical studies in interaction design**.

Keywords

Human-Robot Interaction, Framing, Perceived Trustworthiness, Trust Calibration, Overtrust, Mental Models, Industrial Robots

1 Introduction

Even in highly automated environments, such as in the food industry, human operators retain a supervisory role for oversight on different levels, including intervention, and possible error recovery [30]. The ability of the human operator to maintain an appropriate level of reliance on the system - neither over-relying on the automation nor under-utilizing it - directly affects the safety and efficiency of system performance [15, 18].

Historically referred to as a trust calibration (i.e., [3]), this process relies on various factors, such as individual human traits, context of the interaction, aspects of the system itself (i.e., [8]), and how the system capabilities are conveyed (i.e., [11]). Transparency is a natural means for trust calibration [15]. In this paper, we use transparency as an umbrella term for information that helps users in trustworthiness assessment: to understand what the system is doing, why it is doing it, and what to expect next [16]. This information can be provided not only during interaction, but also as a priori displacement before the first use [19] (i.e., during onboarding or by providing documentation or training). For this work, trust is one aspect constitutive of the human-robot relation in the workplace. We understand it as the willingness of the user to rely on the robot's action [18] under the condition of uncertainty [15], based on the expectations they form about the system's capabilities. Various studies examined how people form trust and reliance in different application domains [16, 26, 28], however, much less is known about how people form trusting behavior of the system based on early trust cues in industrial settings (i.e., overview in [2],[9]).

Research on how capabilities of the system are presented to the user shows that early descriptions of the system's capabilities can shape users' mental models and trust judgments [9, 20, 23, 29]. Related work in HRI and pragmatics has similarly shown that the linguistic introduction and description of a robot have an effect on users' expectations of what the robot can do [6]. In HRI, framing has been established as a concept to capture the ways in which a robot and its capabilities to act and interact are presented to (future) users and operators by the developers (engineers). It refers to providing "structure[s] of expectation" [31]— information that shapes how users interpret its purpose, or capabilities prior to the interaction [22]. The term framing thus describes a practice of introducing a robot as a capable interlocutor. At the same time, it is an analytical term, used to explore and understand framing as a practice that grants a robot more than purely technical, operational functions, but rather embeds it in a relation of seemingly reciprocal expectations.

Previous work on framing focused on communicated privacy, or social ability [27]; anthropomorphism vs functional design [22], and [23] showed that verbal expectation setting can influence users' perspectives on robot's competences. The findings show mixed results in their experiments with personal robots. Kopp et al. show that linguistic framing changes the student's trust in a collaborative robot, and that the effect depends on whether the robotic arm is framed as a cooperative partner Paul, versus a machine [13], however, at the same time, they highlighted the need for context-sensitivity and dependence on individual characteristics of the recipients of the framing. Most closely related, Washburn et al. showed that functionality framing shapes perceived reliability and trustworthiness in a collaborative HRI task with a robot that makes intentional errors [34]. This highlights that early transparency about the robot functionality and capability is not neutral; it shapes initial trustworthiness assessments, even before users can observe meaningful performance evidence.

In industrial settings, HRI research typically focuses on objective measures, such as performance metrics, ergonomics, safety, and alike [14, 33]. However, little is known about subjective trustworthiness and how the subjective trustworthiness of the system affects the behavioral outcomes of the interaction. Prior work shows that early cues can bias users toward either unwarranted caution or over-reliance [5, 24]. Such biases affect operator attention, intervention timing, and overall human-automation team performance [8, 15]. Building on this, we expect (or hypothesize) that framing might shape perceived trustworthiness, even when an actual system's performance does not (yet) justify these perceptions. However, we still lack empirical evidence on how these early cues function when the robot is neither a social actor nor possesses anthropomorphic cues [22]. This is particularly relevant in industrial scenarios, where operators often have to form rapid judgments about system capability without social and anthropomorphic cues available in non-industrial settings.

This article examines the relationship between system descriptions and oral explanations as initial framing cues on the one hand, and the emerging perceived trustworthiness of the system on the other. We centrally ask how communicated competence, that is, framing the system as either a low-capability prototype or a high-capability advanced technology, shapes the perceived trustworthiness of the system.

Our paper focuses on a human-robot teaming with an industrial food-packing robot. We designed a case study in which we kept the system behavior constant and varied only the framing. Through this, our study aimed to isolate the influence of cues on trust formation that were communicated before any form of actual interaction, with a focus on perceived trustworthiness.

We aim to contribute empirical evidence to the effects of framing in HRI. We present evidence that a priori transparency about capability [19] should be treated as a trustworthiness cue that can influence operators' reliance, and therefore must be treated carefully so that operators can later confirm (or reject) them during the interaction with the robot [11, 15].

A virtual reality (VR) setup was used to create a representation of an automated food-packing robot that communicates verbally with the human operator, who collaborates with the robot to ensure good outcome. VR enabled us to experiment in a realistic environment to test the different approaches to the framing of a real robotic system, MOZART, which is currently under development. Furthermore, the setting allowed us to test different trustworthiness cues by providing controlled exposure to the same system behavior while manipulating only the communicated competence (i.e., isolating framing while holding observable performance of the robot constant).

With this rationale, we tested the following two hypotheses:

- H1: Framing the system as high-capability will increase the participants' perceived trustworthiness of the robot over the low-capability framing.
- H2: Participants in the high-capability framing condition will demonstrate more trusting behavior than participants in the low-capability framing condition.

2 Method

2.1 Participants

Twenty-one participants (11 in the low-capability; 10 in the high-capability condition) were recruited in Odense, at the University of Southern Denmark. The sample included 12 men and 9 women with ages ranging from 19 to 51 years (M = 28.48, SD = 8.01). Participants reported low prior experience with the technology used: 57% reported no prior experience with robots, and 52% reported this was their first VR experience. The study followed institutional ethical guidelines and was previously cleared by the ethical board of TCD. All participants in this between-subject study volunteered and received no compensation. They all provided informed consent about their participation, and were informed about the risks linked with using virtual reality, including temporary disorientation or motion sickness. Here, we followed the recommendation for conducting ethical experiments in augmented reality by Madary and Metzinger [17]. No participants reported any discomfort or sickness during or after the VR experiment.

2.2 Scenario in virtual reality

The study was conducted using a Meta Quest Pro headset. The headset displays a VR projection of an automated manipulation and packing system as it is developed for the project MOZART [21]. The VR solution itself was developed at Brno University of Technology based on the work presented in [12]. Participants could observe a scene displaying a Z-shaped industrial line handling and packing chicken drumsticks into a plastic tray.

Figure 1: Simplified AI-generated illustration of the VR scenario. A participant wearing a VR headset supervises the Z-shaped packing line. Chicken drumsticks move onto the robotic tile array (black tiles), which performs orientation and transfer before the line continues to another section. The operator (study subject) controls the system by clapping; the system interacts with the user using voice utterances.

Participants could walk around the virtual packing line and observe the system from different angles. All visual, auditory, and behavioral aspects of the experience were identical across conditions. Background audio included ambient industrial noise (freesound.org, ID 641039). The system was controlled using a Wizard-of-Oz setup [1], a common practice in early-stage HRI experiments, in which the experimenter controlled the audio prompts. For consistency, the same person controlled the system for all participants. Participants spent approximately 80 seconds with the VR headset on.

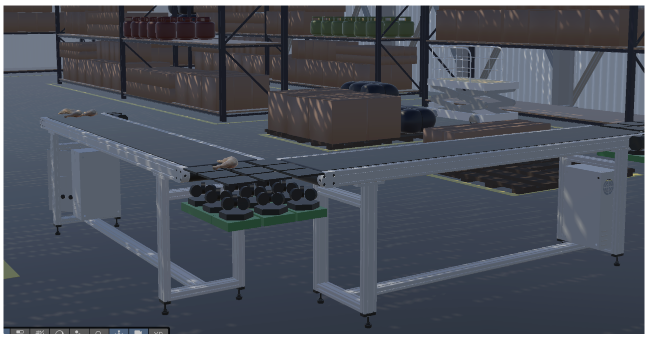

Figure 2: A screenshot of the robotic industrial line, which operators had to control.

2.3 Task

Participants acted as operators responsible for monitoring the poultry meat-packing robot. Specifically, they were responsible for making decisions about the speed of the line. Prior to the interaction with the system, participants received instructions from the experiment facilitators. They were told that the system could make errors during the warm-up and that their job was to:

- Stop the line in case they see any errors, by clapping their hands once.
- Increase the speed of the line by clapping their hands twice.

To appropriately control the system's performance, participants were provided with information about its function and the possible consequences of their actions. They were informed that speeding up the line too early could risk system damage, while delaying the change of speed would increase food waste. They were specifically instructed that it is their responsibility to determine the appropriate moment to increase the speed. The rationale behind this narrative framing of the robot is that, with this, we intended to increase the perceived risk in the task, while providing only a limited, and controlled amount of information about the system.

2.3.1 Experimental Manipulation. As our primary interest was to test the effect of the system's framing on perceived trustworthiness, we created two experimental conditions that differed in how the system's capability was communicated prior to the interaction:

Condition A: Low-capability framing.

> "In this study, you will be overseeing an early prototype of an automated packing system. This version was developed as part of a student research project, exploring whether a low-cost solution could handle chicken drumsticks on a conveyor line. The system is still in an experimental phase, and its performance has only been tested in a small number of lab trials."

Condition B: High-capability framing

> "In this study, you will be overseeing an advanced automated packing system. This version was developed by a team of expert researchers and industry partners working in robotics and AI for food processing. The system uses state-of-the-art sensing and control, and its performance has been extensively tested in a wide range of lab trials."

Importantly, the manipulation focused only on the communicated technical maturity and system validation; the system's behavior remained the same under both experimental conditions.

2.4 Measures

Prior to the start of each experiment, participants completed a basic demographic questionnaire in line with the standards commonly required by the HRI community [35]. Using a tablet, we also collected data on their prior experience with robots, virtual reality, and their propensity to trust, based on [7].

To capture the self-reported level of perceived trustworthiness, we utilized the Multidimensional Measure of Trust (MDMT). More specifically, we used the subscales for Reliability and Competence of the MDMT v2 (2020-09-01) [32]. Participants rated each item on a 0-7 Likert scale (0 = not at all, 7 = very). As required by the instructions, we also included a "does not fit" option, treated as a missing value in the subsequent analysis.

Furthermore, we included three questions designed to provide us with insights into whether the test subjects were affected by the framing in the two experimental conditions. We asked about how i) professionally engineered, ii) validated, and iii) extensively tested the system was. We also collected data about how people interacted with the system: the behavioral measure was the latency (in seconds) between the end of the first instructions given by the robot and the participant's first "clap twice" action.

2.5 Procedure

Participants were approached in the corridor at the University of Southern Denmark, and asked whether they had 10-15 minutes to voluntarily participate in a short VR study. After agreeing, they were guided to the research laboratory, given a tablet with the consent form and initial survey. Participants were then seated, received a short task introduction, including the framing manipulation, and were invited to ask clarifying questions. Once ready, the experimenter assisted them in putting on the headset. The virtual environment initially displayed only a warehouse, allowing participants to familiarize themselves with it. The ambient factory noise began to play, after which the experimenter read one of the two framing scripts and asked participants to stand up and look around for the production line.

Next, a pre-recorded audio script started. After ten seconds of the ambient noise, participants heard the robotic system verbally present a brief task explanation, instructing them to monitor the line and decide when to speed it up or stop it by clapping their hands. When this utterance ended, the experimenter started a stopwatch, which ran until participants interacted with the system by clapping their hands twice. This double clap prompted a pre-recorded response from the system ("Request for a change of speed detected ...", followed 20 seconds later by a second message ("Final speed achieved..."). Participants were then informed that the experiment was concluded and were asked to remove the headset.

Two participants (one from each condition) performed a single clap during the experiment, which triggered a corresponding audio prompt asking them to confirm a stop request or not take any further action. Neither of them confirmed the request to stop the line, and both subsequently clapped twice and successfully completed the experiment at their own pace and their total time was counted.

Finally, participants completed the remaining questionnaire items on perceived trustworthiness and the manipulation check. A short debriefing followed, informing participants that the system was controlled by the facilitator and informing them about the objectives of the experiment.

Figure 3: A participant wearing the Meta Quest Pro headset during the VR task

3 Results

Data were collected via SurveyXact [25], and imported to SPSS 30 [10] for additional analysis. Missing values (one participant - using "does not apply" option of the MDMT) were treated in the data file. In addition, in line with the instructions of [32], two average values were calculated for the Reliability and Competence sub-scales of the MDMT scale. Independent samples t-tests were conducted to evaluate the effects of framing on perceived trustworthiness and behavior.

3.1 Manipulation Check

Participants in the high-capability framing condition rated the system as more professionally engineered, more extensively validated, and more thoroughly tested than those in the low-capability condition. All effects were statistically significant (see Table 1), which confirms that the framing manipulation was successful.

3.2 Perceived Trustworthiness

Perceived trustworthiness of the system was measured using the Reliability and Competence sub-scales of the MDMT[32]. Participants presented with the high-capability framing rated the system as more competent (t(18) = −2.39, p = .029, d = 1.59) and more reliable (t(18) = −2.00, p = .061, d = 1.43) compared to those exposed to the system framed as the low-capability one. The effect for Reliability did not reach conventional significance with a two-tailed test. Thus, H1 was supported for Competence, while Reliability showed an effect in the hypothesized direction.

3.3 Behavioral Measure

The experiment also focused on quantifying the trusting behavior of participants, which was operationalized in the time participants took before they decided to increase the speed of the line. Participants in the high-capability condition reacted faster (M = 21.0 s, SD = 11.3) than those in the low-capability condition (M = 44.9 s, SD = 39.2). This difference in means is consistent with H2: participants in the high-capability framing appeared to show more trusting behavior toward the system, even though they interacted with the same robotic system. However, the results did not reach the conventional level of significance using the two-tailed test (t(19) = 1.86, p = .079), so this result should be rather interpreted as a directional trend.

Table 1: Descriptive statistics and independent-samples *t*-tests for trustworthiness measures, behavioral response time, and manipulation-check items.

Measure	Cond.	N	M	SD	*t*	df	*p*	d
MDMT Reliability	Low	10	5.03	1.64	-2.00	18	.061	1.43
	High	10	6.25	1.03				
MDMT Competence	Low	10	4.50	1.74	-2.39	18	.029	1.59
	High	10	6.13	1.27				
Reaction Time (s)	Low	11	44.91	39.18	1.86	19	.079	0.81
	High	10	21.00	11.29				
Perc. Professional	Low	11	3.82	1.60	-3.16	19	.005	1.44
	High	10	5.80	1.23				
Perc. Validation	Low	11	3.18	1.54	-2.75	19	.013	1.60
	High	10	5.10	1.66				
Perc. Testing	Low	11	4.18	1.66	-2.29	19	.033	1.42
	High	10	5.60	1.08				

Note. Abbreviated measure labels are used for space reasons: Reaction Time = reaction time to speed increase; Perc. Professional = perceived professional engineering; Perc. Validation = perceived validation level; Perc. Testing = perceived testing extensiveness. MDMT analyses exclude one participant who selected "does not fit."

We additionally analyzed whether the participant's individual propensity to trust [7] predicted the perceived trustworthiness of the system or the effect of framing. Propensity to trust is a general willingness to trust others. It influences initial trust judgments in situations with limited information [18]. Pearson correlations showed that in our sample, trust propensity was not significantly associated with MDMT Reliability (r = .25, p = .30) or MDMT Competence (r = .24, p = .31).

4 Discussion

This study examined how a priori transparency and communicated competence—verbal framing of an industrial robot system—shapes perceived trustworthiness and trusting behavior. Our results showed the same directional pattern across the self-reported and behavioral measures, although not all effects reached the conventional level of significance. Participants exposed to the system framed as "more competent" perceived the system as more competent and reliable on the MDMT subscales. The ratings differed despite participants in the low-capability condition observing the same system's behavior. This extends previous work on framing, showing that early cues affect user expectations [5, 24], even when the system's actual behavior remains unchanged.

The effect of framing manipulation showed the same directional pattern. Participants in the high-capability condition prompted the system to speed up sooner: on average, they waited for 21 seconds, compared with 45 seconds in the low-capability framed condition. Although not statistically significant in the two-tailed test, the results are consistent with the data on perceived trustworthiness; the lower statistical power may be due to the small sample size. Still, the pattern is consistent with the interpretation that communicated competence is linked to perceived trustworthiness, which guides decision-making in situations where system behavior is ambiguous or has not yet been observed. In this regard, our findings resonate with work demonstrating that framing can outweigh observed robot behavior in early interaction phases [23].

An essential aspect of our study was that participants interacted with a short, immersive VR scenario in which behavioral cues were limited. Under such conditions, framing may serve as the primary interpretative cue, especially since the operator's role was linked to a perceived risk (i.e., damaging the line or wasting food). This may explain why the effect on perceived trustworthiness was more substantial than in studies focusing on anthropomorphic framing [22]. The results indicate that when early system behavior provides insufficient evidence of competence, users may rely on the framing as the most reliable cue.

This suggests that, in the early stages of the industrial HRI, when observable behavior is limited or ambiguous, inappropriate framing can contribute to miscalibrated trust. If the way how the robot is framed overstates the system's capability, operators may prematurely rely on it; if framing understates the system's capability, operators may underutilize it. From a trust-calibration perspective, this study confirms that framing can act as a high-impact transparency cue: it can raise or dampen reliance before operators have sufficient evidence to evaluate and validate the system's real capabilities and limits [4, 15]. This creates a practical implication for industrial HRI: early capability descriptions should be designed so they can be verified (or rejected/corrected) during use, and so they do not overstate their capabilities in ways that promote premature reliance [11]. Future work is needed to examine long-term interactions with more fine-grained levels of framing and how these contribute to trust development and future trust calibration.

5 Conclusion

In this paper, we explored the relationship between a communicative framing of a robot prior to interaction and its impact on early trust formation. As we show, the perceived trustworthiness is highly dependent on the communicative framing: participants exposed to the high-capability framing rated the system as more competent, showed higher reliability ratings, and tended to act more quickly, despite identical system performance (see Table 1). These findings are in line with prior work showing that early setting of expectation through framing (i.e., related to functionality) can shape the level of perceived trustworthiness and responses to robot behavior [23, 34]. Taken together, this suggests that framing in industrial scenarios should be treated as a strategic design lever to foster the ideal of calibrated trust in HRI: capability descriptions should be chosen deliberately to support appropriate reliance, and should be paired with opportunities for operators to verify, reject, or revise these expectations not only in the early stages of interaction [4, 11, 15].

We are aware of the limitations of this case study: the small sample size, the exclusively quantitative data, the specific, short mediated interaction through VR, and the absence of a baseline condition. Also, the perceived risk was only introduced narratively without any real behavioral consequences. Nevertheless, our findings illustrate that participants relied strongly on the communicated framing to form expectations, outweighing the (limited) behavioral evidence available from their short interaction. This suggests a risk of miscalibration in the early stages of industrial HRI, when system capabilities are overstated or underspecified. Finally, we highlight the importance of framing and emphasize the need to address it further in future research, as it plays a significant role in forming appropriate reliance not only in early (industrial) HRI.

6 Acknowledgments & Funding

The authors thank the engineering team at Brno University of Technology, especially Michal Kapinus, for the help with implementing the VR scenario. The work on this article was funded by the EU-HORIZON RIA project MOZART (Morphing Computerized mats with Embodied Sensing and Artificial Intelligence), grant agreement no 101069536).

References

[1] Christel Brin{\textbackslash}-döpke, Michaela Johanntokrax, Arno Pahde, and Britta Wrede. 1995. Darf ich Dich Marvin nennen? Instruktionsdialoge in einem Wizard-of-Oz Szenario: Materialband.

[2] Giulio Campagna and Matthias Rehm. 2024. A Systematic Review of Trust Assessments in Human-Robot Interaction. (2024), 3706123. doi:10.1145/3706123

[3] Ewart de Visser, Marvin Cohen, Amos Freedy, and Raja Parasuraman. 2014. *A Design Methodology for Trust Cue Calibration in Cognitive Agents*. Vol. 8525. doi:10.1007/978-3-319-07458-0_24 Pages: 262.

[4] Ewart J. de Visser, Marieke M. M. Peeters, Malte F. Jung, Spencer Kohn, Tyler H. Shaw, Richard Pak, and Mark A. Neerincx. 2020. Towards a Theory of Longitudinal Trust Calibration in Human–Robot Teams. *International Journal of Social Robotics* 12, 2 (2020), 459–478. doi:10.1007/s12369-019-00596-x

[5] Mary T. Dzindolet, Scott A. Peterson, Regina A. Pomranky, Linda G. Pierce, and Hall P. Beck. 2003. The role of trust in automation reliance. 58, 6 (2003), 697–718. doi:10.1016/S1071-5819(03)00038-7

[6] Kerstin Fischer. 2011. How People Talk with Robots: Designing Dialog to Reduce User Uncertainty. 32, 4 (2011), 31–38. doi:10.1609/aimag.v32i4.2377

[7] M. Frazier, Paul Johnson, and Stav Fainshmidt. 2013. Development and Validation of a Propensity to Trust Scale. 3 (2013), 76–97. doi:10.1080/21515581.2013.820026

[8] Peter A. Hancock, Deborah R. Billings, Kristin E. Schaefer, Jessie Y. C. Chen, Ewart J. de Visser, and Raja Parasuraman. 2011. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. 53, 5 (2011), 517–527. doi:10.1177/0018720811417254

[9] Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in automation: Integrating empirical evidence on factors that influence trust. 57, 3 (2015), 407–434. Publisher: Sage Publications Sage CA: Los Angeles, CA.

[10] IBM Corp. 2023. *IBM SPSS Statistics for Windows, Version 30.0.* Armonk, NY.

[11] Matouš Jelínek, Nadine Schlicker, and Ewart de Visser. 2025. Six Guidelines for Trustworthy, Ethical and Responsible Automation Design. *arXiv preprint arXiv:2508.02371* (2025).

[12] Michal Kapinus, Vítězslav Beran, Zdeněk Materna, and Daniel Bambušek. 2024. Augmented Reality Spatial Programming Paradigm Applied to End-User Robot Programming. *ROBOTICS AND COMPUTER-INTEGRATED MANUFACTURING*

89, 89 (2024), 1–13. doi:10.1016/j.rcim.2024.102770

[13] Tobias Kopp, Marco Baumgartner, and Steffen Kinkel. 2023. "It's not Paul, it's a robot": The impact of linguistic framing and the evolution of trust and distrust in a collaborative robot during a human-robot interaction. *International Journal of Human-Computer Studies* 178 (2023), 103095.

[14] Przemyslaw A Lasota, Terrence Fong, Julie A Shah, et al. 2017. A survey of methods for safe human-robot interaction. *Foundations and Trends® in Robotics* 5, 4 (2017), 261–349.

[15] John D. Lee and Katrina A. See. 2004. Trust in Automation: Designing for Appropriate Reliance. 46, 1 (2004), 50–80. doi:10.1518/hfes.46.1.50_30392

[16] Joseph B Lyons. 2013. Being Transparent about Transparency: A Model for Human-Robot Interaction. (2013), 6.

[17] Michael Madary and Thomas K Metzinger. 2016. Real virtuality: A code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology. *Frontiers in Robotics and AI* 3 (2016), 180932.

[18] Roger C. Mayer, James H. Davis, and F. David Schoorman. 1995. An Integrative Model of Organizational Trust. 20, 3 (1995), 709–734. doi:10.2307/258792

[19] Christopher A Miller. 2021. Trust, transparency, explanation, and planning: Why we need a lifecycle perspective on human-automation interaction. In *Trust in human-robot interaction*. Elsevier, 233–257.

[20] Scott Anthony Mishler. 2023. *Framing Automation Trust: How Initial Information About Automated Driving Systems Influences Swift Trust in Automation and Trust Repair for Human Automation Collaboration*. Ph. D. Dissertation. Old Dominion University.

[21] MOZART Consortium. 2026. MOZART Project. https://mozart-robotics.eu/project/ Accessed 2026-03-06.

[22] Linda Onnasch and Eileen Roesler. 2019. Anthropomorphizing robots: The effect of framing in human-robot collaboration. In *Proceedings of the human factors and ergonomics Society annual meeting*, Vol. 63. SAGE Publications Sage CA: Los Angeles, CA, 1311–1315.

[23] Steffi Paepcke and Leila Takayama. 2010. Judging a bot by its cover: An experiment on expectation setting for personal robots. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 45–52.

[24] Raja Parasuraman and Victor Riley. 1997. Humans and Automation: Use, Misuse, Disuse, Abuse. 39, 2 (1997), 230–253. doi:10.1518/001872097778543886 Publisher: SAGE Publications Inc.

[25] Rambøll Management Consulting. 2023. *SurveyXact*. Copenhagen, Denmark.

[26] Paul Robinette, Wenchen Li, Robert Allen, Ayanna M. Howard, and Alan R. Wagner. 2016. Overtrust of robots in emergency evacuation scenarios. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Christchurch, New Zealand, 2016-03). IEEE, 101–108. doi:10.1109/HRI.2016.7451740

[27] Matthew Rueben, Frank J Bernieri, Cindy M Grimm, and William D Smart. 2017. Framing effects on privacy concerns about a home telepresence robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 435–444.

[28] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. *Would You Trust a (Faulty) Robot?: Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust*. Vol. 2015. doi:10.1145/2696454.2696497

[29] Nadine Schlicker, Kevin Baum, Alarith Uhde, Sarah Sterz, Martin C Hirsch, and Markus Langer. 2025. How do we assess the trustworthiness of AI? Introducing the trustworthiness assessment model (TrAM). *Computers in Human Behavior* 170 (2025), 108671.

[30] Thomas Sheridan. 2016. Human-Robot Interaction: Status and Challenges. 58 (2016). doi:10.1177/0018720816644364

[31] Deborah Tannen. 1977. What's in a frame? Surface evidence for underlying expectations. *New direction in discourse processing* (1977).

[32] Daniel Ullman and Bertram F. Malle. 2019. Measuring Gains and Losses in Human-Robot Trust: Evidence for Differentiable Components of Trust. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2019-03). IEEE, 618–619. doi:10.1109/HRI.2019.8673154 Place: Daegu, Korea (South).

[33] Valeria Villani, Fabio Pini, Francesco Leali, and Cristian Secchi. 2018. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics* 55 (2018), 248–266.

[34] Auriel Washburn, Akanimoh Adeleye, Thomas An, and Laurel D Riek. 2020. Robot errors in proximate hri: how functionality framing affects perceived reliability and trust. *ACM Transactions on Human-Robot Interaction (THRI)* 9, 3 (2020), 1–21.

[35] Katie Winkle, Erik Lagerstedt, Ilaria Torre, and Anna Offenwanger. 2023. 15 years of (who) man robot interaction: Reviewing the h in human-robot interaction. *ACM Transactions on Human-Robot Interaction* 12, 3 (2023), 1–28.

Received 20 February 2026