System Transparency in Shared Autonomy: A Mini Review

Victoria Alonso[1] and Paloma de la Puente[2,3]*

[1] ETSI Aeronáutica y del Espacio, Universidad Politécnica de Madrid, Madrid, Spain, [2] ETSI Industriales, Universidad Politécnica de Madrid, Madrid, Spain, [3] Centre for Automation and Robotics, Universidad Politécnica de Madrid-CSIC, Madrid, Spain

What does transparency mean in a shared autonomy framework? Different ways of understanding system transparency in human-robot interaction can be found in the state of the art. In one of the most common interpretations of the term, transparency is the observability and predictability of the system behavior, the understanding of what the system is doing, why, and what it will do next. Since the main methods to improve this kind of transparency are based on interface design and training, transparency is usually considered a property of such interfaces, while natural language explanations are a popular way to achieve transparent interfaces. Mechanical transparency is the robot capacity to follow human movements without human-perceptible resistive forces. Transparency improves system performance, helping reduce human errors, and builds trust in the system. One of the principles of user-centered design is to keep the user aware of the state of the system: a transparent design is a user-centered design. This article presents a review of the definitions and methods to improve transparency for applications with different interaction requirements and autonomy degrees, in order to clarify the role of transparency in shared autonomy, as well as to identify research gaps and potential future developments.

Keywords: transparency, shared autonomy, human-robot interaction, communication, observability, predictability, interface, user-centered design

1. INTRODUCTION

Shared autonomy adds to the fully autonomous behavior some level of human interaction, combining the strengths of humans and automation (Hertkorn, 2015; Schilling et al., 2016; Ezeh et al., 2017; Nikolaidis et al., 2017). In shared autonomy, humans and robots have to collaborate. Transparency supports a flexible and efficient collaboration and plays a role of utmost importance regarding the system overall performance.

In the next sections, current research about transparency in the shared autonomy framework is reviewed. The goal is to provide, by analyzing the literature, a general view for a deeper understanding of transparency which helps motivate and inspire future developments. The key aspects and most relevant previous findings will be highlighted.

Different ways of understanding transparency in human-robot interaction in the shared autonomy framework can be found in the state of the art. In one of the most common interpretations of the term, transparency is the observability and predictability of the system behavior, the understanding of what the system is doing, why, and what it will do next.

In section 2 the effect of levels of autonomy on transparency is analyzed. Then, the mini-review is organized according to the different ways of understanding transparency in human-robot interaction in the shared autonomy framework.

In section 3 transparency as observability and predictability of the system behavior is studied. Since the main methods to improve transparency are based on interface design and training, transparency is usually considered a property of such interfaces, and section 4 focuses on transparency as a property of the interface. Since natural language explanations are a popular way to achieve transparent interfaces, transparency as explainability is studied in section 5. Section 6 is dedicated to mechanical transparency, and ethically aligned design aspects of transparency are reviewed in section 7.

Hence, the wider and most extended interpretations and results are presented first, while more specific trends are left for later sections. This way, the reader can naturally focus on the general concepts before other implications are analyzed. A table of selected references for each section can be found at the end of the paper (**Table 1**).

2. TRANSPARENCY AND LEVELS OF AUTONOMY

Traditionally, human-robot interaction in shared autonomy has been characterized by levels of autonomy (Beer et al., 2014). Sheridan and Verplank (1978) proposed an early scale of levels of autonomy to provide a vocabulary for the state of interaction during the National Aeronautics and Space Administration (NASA) missions. In later work, Endsley and Kaber (1999) and Parasuraman et al. (2000) established other levels of autonomy taxonomies, considering the distribution of tasks between the human and the system regarding information acquisition, decision making, and actions implementation.

Recently, Kaber (2017) reopened the discussion about whether levels of autonomy are really useful. This paper has received answers from Miller (2017) and Endsley (2018). Miller (2017) considers that levels of autonomy are only an attempt to reduce the dimension of the multidimensional space of human-robot interaction. Other authors agree with this multidimensional perspective (Bradshaw et al., 2013; Gransche et al., 2014; Schilling et al., 2016) and with the need of focusing on human-robot interaction (DoD, 2012). Endsley's reply (Endsley, 2018) is a review about the benefits of levels of autonomy.

For high levels of autonomy, when the system is operating without significant human intervention, additional uncertainty is expected. For high levels of autonomy, the user may have a low level of observability of the system behavior, and low predictability of the state of the system. The system might have a low level of transparency.

For low levels of autonomy, when the human operators are doing almost everything directly themselves, they know how the tasks are being carried out, so the uncertainty and unpredictability are typically low. Yet, the human operator's cognitive workload to be aware of everything increases to process all the information. If the cognitive workload is too high, a solution is delegation (Miller, 2014). Without trust, the user is not going to delegate, no matter how capable the robot is, under-trust may cause a poor use of the system (Parasuraman and Riley, 1997; Lee and See, 2004). Transparency is needed for understanding and trust, and trust is necessary for delegation (Kruijff et al., 2014; Ososky et al., 2014; Yang et al., 2017).

The cognitive workload reduction not always means a task performance improvement because of the automation-induced complacency (Wright et al., 2017). Complacency means over-trusting the system, and it is defined in Parasuraman et al. (1993) as "the operator failing to detect a failure in the automated control of system monitoring task." For high levels of autonomy, transparency of the robot intent and reasoning is especially necessary to make the most of human-in-the-loop approaches, reducing complacency (Wright et al., 2017). Transparency and trust calibration can be improved by training (Nikolaidis et al., 2015) and a good interface design (Lyons, 2013; Kruijff et al., 2014). Some efforts to integrate trust into computational models can be found in Desai (2012) and Chen et al. (2018).

3. TRANSPARENCY AS OBSERVABILITY AND PREDICTABILITY OF THE SYSTEM BEHAVIOR

One of the most common ways of understanding transparency in human-robot interaction in shared autonomy framework is as observability and predictability of the system behavior: the understanding of what the system is doing, why, and what it will do next (Endsley, 2017).

What kind of information should be communicated in order to have a good level of transparency? The robot's state and capabilities must be communicated transparently to the human operator: what the robot is doing and why, what it is going to do next, when and why the robot fails when performing specific actions, and how to correct errors are essential aspects to be considered. In Kruijff et al. (2014) and Hellström and Bensch (2018), the authors go even further: their research explores, based on experimental data, not only what to communicate, but also communication patterns—how to communicate—for improving user understanding in a given situation.

Autonomy increases uncertainty and unpredictability about the system's state, and some authors understand transparency in the sense of predictability: "Transparency is essentially the opposite of unpredictability" (Miller, 2014) and "Transparency is the possibility to anticipate imminent actions by the autonomous system based on previous experience and current interaction" (Iden, 2017).

Other definitions, found in the literature, in the sense of observability are: "Transparency is the term used to describe the extent to which the robot's ability, intent, and situational constraints are understood by users" (Wortham et al., 2017), "Transparency is a mechanism to expose the decision-making of a robot" (Theodorou et al., 2016, 2017), and "the ability for the automation to be inspectable or viewable in the sense that its mechanisms and rationale can be readily known" (Miller, 2018).

TABLE 1 | Summary.

Transparency in shared autonomy			
2. Transparency and levels of autonomy			
High autonomy		Low autonomy	Miller, 2014, 2017
Situation awareness	Low	High	Schilling et al., 2016; Ezeh et al., 2017
Transparency	Low	High	Kaber, 2017; Wright et al., 2017
Cognitive engagement	Low	High	Ososky et al., 2014; Yang et al., 2017
Risk of overtrust	High	Low	Desai, 2012; Chen et al., 2018
3. Transparency as observability and predictability of system behavior			
Transparency as observability of the system behavior.			Endsley, 2012, 2017
User understanding of what the system is doing, why, and what it will do next. User-centered design.			Kruijff et al., 2014; Hellström and Bensch, 2018; Villani et al., 2018
Transparency as the opposite of unpredictability			Miller, 2014; Iden, 2017
The robot's abilities, intent, and situational constraints are understood by the users. Legibility, Readibility			Takayama et al., 2011; Dragan et al., 2013; Busch et al., 2017; Wortham et al., 2017
Mechanism to expose decision making			Theodorou et al., 2017
Methods to establish Shared Situational Awareness and Shared Intentions (Interface Design and Training) Robot-to-Human and Robot-of-Human Transparency			Lyons, 2013; Lyons and Havig, 2014; Tsiourti and Weiss, 2014; Lorenz, 2015; Dragan, 2017; Wang et al., 2017; Doellinger et al., 2018; Javdani et al., 2018
4. Transparency as a property of the Interface			
Situation Awareness Transparency (SAT) Model.			Chen et al., 2014; Endsley, 2017
Levels 1,2,3: Perception, comprehension, and projection			
Multimodal interfaces			Perzanowski et al., 2001; Lakhmani et al., 2016; Oviatt et al., 2017
Transparency as explainability			
Transparency is the robot offering explanations of its actions. Route/planning/navigation verbalization			Kim and Hinds, 2006; Caminada et al., 2014; DARPA, 2016; Rosenthal et al., 2016; Wortham and Rogers, 2017
5. Mechanical transparency			
Wearable robots,exoskeletons,rehabilitation: Capacity to follow human movements without human-perceptible resistive forces Training, Avoiding over-trust			Robertson et al., 2007; Jarrasse et al., 2008; Jarrassé et al., 2009; van Dijk et al., 2013; Zhang et al., 2016; Awad et al., 2017; Beckerle et al., 2017; Bai et al., 2018; Borenstein et al., 2018; Fani et al., 2018
Telerobotics Realistic (transparent) perception of the remote environment.			Raju et al., 1989; Lawrence, 1993; Yokokohji and Yoshikawa, 1994; Ferre et al., 2007; Hirche and Buss, 2007; Slawinski et al., 2012; Goodrich et al., 2013; Hertkorn, 2015; Muelling et al., 2017
6. Transparency and ethically aligned design			
The ethical black-box			Winfield and Jirotka, 2017
Transparency as traceability and verification			Wortham et al., 2017
IEEE Global Initiative for Ethically Aligned Design			Bryson and Winfield, 2017
P7001-Transparency in Autonomous Systems			Grinbaum et al., 2017

Legible motion is "the motion that communicates its intents to a human observer" (Dragan et al., 2013), also referred as readable motion (Takayama et al., 2011), or anticipatory motion (Gielniak and Thomaz, 2011). Algorithmic approaches for establishing transparency in the sense of legibility can be found in Dragan et al. (2013, 2015a,b); Nikolaidis et al. (2016), and in Busch et al. (2017) based on optimization and learning techniques, respectively.

3.1. Robot-to-Human Transparency and Robot-of-Human Transparency

Transparency about the robot's state information may be referred to as robot-to-human transparency (Lyons, 2013). One of the principles of user-centered design is to keep the user aware of the state of the system (Endsley, 2012; Villani et al., 2018). Robot-to-human transparency enables user-centered design. This mini-review is focused on this type of transparency.

There is also a robot-of-human transparency (Lyons, 2013), which focuses on the awareness and understanding of information related to humans. This concept of monitoring human performance is of growing interest to provide assistance, e.g., in driving and aviation. The term robot-of-human transparency is not widely used in literature. However, examples of robot-of-human transparency, without using the term directly, can be found in Lorenz et al. (2014); Lorenz (2015); Tsiourti and Weiss (2014); Dragan (2017); Wang et al. (2017); Doellinger et al. (2018); Goldhoorn et al. (2018); Gui et al. (2018), and Javdani et al. (2018). In Casalino et al. (2018) and Chang et al. (2018) a feedback of the intent recognition is communicated to the operator.

In Lyons (2013) and Lyons and Havig (2014) transparency is defined as a "method to establish shared intent and shared awareness between a human and a machine." Since the main method to establish shared situation awareness and shared intent is the interface design, the next section is dedicated to the study of transparency as a property of the interface.

4. TRANSPARENCY AS A PROPERTY OF THE INTERFACE

The Human-Automation System Oversight (HASO) model (Endsley, 2017) summarizes the main aspects, and its relationships, of Human-Automation Interaction (HAI). The place of transparency in this model is as a property of the interface. This model uses the three level situation awareness model (Endsley, 1995).

In Chen et al. (2014) Transparency is defined as an attribute of the human-robot interface "the descriptive quality of an interface about its abilities to afford an operator's comprehension about an intelligent agent's intent, performance, plans, and reasoning process." The Situation Awareness Transparency (SAT) model (Chen et al., 2014), is based on Endsley (1995), and proposes three levels of Transparency:

- Level 1. Transparency to support perception of the current state, goals, planning, and progress.
- Level 2. Transparency to support comprehension of the reasoning behind the robot's behavior and limitations.
- Level 3. Transparency to support projection, predictions and probabilities of failure/success based on the history of performance.

Errors in the perception because the information was not clearly provided (lack of level 1 transparency) are the cause of a great amount of the situation awareness problems, which are the cause of failures due to human errors (Jones and Endsley, 1996; Murphy, 2014). The design of more transparent interfaces might improve situation awareness, reducing human errors.

Information to support transparency can be exchanged through different communication channels (Goodrich and Schultz, 2007): visual interfaces (Baraka et al., 2016; Walker et al., 2018), human-like explanation interfaces (the next sections is dedicated to explanation interfaces), physical interaction and haptics based interfaces (Okamura, 2018) (studied in the mechanical transparency section), or a combination in multimodal interfaces (Perzanowski et al., 2001; Oviatt et al., 2017). Lakhmani et al. (2016) study the possibility to add information about roles and responsibilities in the division of tasks to the SAT model, using a multimodal interface.

5. TRANSPARENCY AS EXPLAINABILITY

Transparency can be achieved by means of human-like natural language explanations. In Kim and Hinds (2006) the definition given for transparency is "Transparency is the robot offering explanations of its actions." Mueller sees explanation as one of the main characteristics of transparency (Mueller, 2016; Wortham et al., 2016).

According to the report about explainable artificial intelligence by the Defense Advanced Research Projects Agency (DARPA, 2016), the explanation interface should be able, at least, to generate answers to the user's questions:

- Why did the system do that and not something else?
- When does the system succeed? and
- When does the system fail?
- When can the user trust the system?
- How can the user correct an error?

Verbalization has been used to convert sensor data into natural language, to describe a route (Perera et al., 2016; Rosenthal et al., 2016) when the user requests information in a dialog, to explain a policy (Hayes and Shah, 2017), or in Zhu et al. (2017) to describe what a humanoid is doing in the kitchen.

Trust in robots is essential for the acceptance and wide utilization of robot systems (Kuipers, 2018; Lewis et al., 2018). Explanations improve usability and let the users understand what is happening, building the users' trust and generating calibrated expectations about the system's capabilities (Westlund and Breazeal, 2016). If systems can explain their reasoning, they should be easily understood by their users, and humans are more likely to trust systems that they understand (Sanders et al., 2014; Sheh, 2017; Fischer et al., 2018; Lewis et al., 2018).

6. MECHANICAL TRANSPARENCY

Wearable robots like exoskeletons are coupled to the user, and the robot moves with the wearer cooperatively (Awad et al., 2017; Anaya et al., 2018; Bai et al., 2018; Fani et al., 2018). In this case, the design should be able to follow the human movements minimizing resistive forces felt by the human, i.e., the design should be mechanically transparent. For example, in rehabilitation, a robot applies a force to a patient, and then the patient finishes the movement (Robertson et al., 2007; Jarrassé et al., 2009; Zhang et al., 2016; Beckerle et al., 2017).

The system is transparent if the robot follows exactly the human movement, without applying forces to the human. Transparency might be improved by human motion prediction (Jarrasse et al., 2008) and training (van Dijk et al., 2013). Trust calibration is needed to avoid the risk of overtrust in the capabilities of the exoskeletons (Borenstein et al., 2018).

Bilateral teleoperation, also named telerobotics, should enable the user to interact with a remote environment as if they were interacting directly. To interact with the remote environment a slave robot is used. The slave is controlled by a human operator using a human-machine interface or master, and the signals from master to slave, and the feedback from slave to master, are transmitted through a communication channel (Ferre et al., 2007; Goodrich et al., 2013; Hertkorn, 2015; Fani et al., 2018; Okamura, 2018).

In a transparent system, the slave tracks exactly the master, and the operator has a realistic (transparent) perception of the remote environment: the technical system should not be felt by the human (Hirche and Buss, 2007). Transparency can be degraded if there are time delays in the communication channel between the user and the remote environment (Lawrence, 1993; Hirche and Buss, 2007; Farooq et al., 2016). More details about transparency modeling for telerobotics can be found in Raju et al. (1989); Lawrence (1993); Yokokohji and Yoshikawa (1994); Ferre et al. (2007), and Slawinski et al. (2012).

When using brain computer interfaces (BCIs) (Bi et al., 2013; Rupp et al., 2014; Arrichiello et al., 2017; Burget et al., 2017) as the input device to teleoperate a robotic manipulator, the difficulty in decoding neural activity introduces delays, noises, etc., and specific techniques to improve transparency are required, such as the ones proposed in Muelling et al. (2017).

7. TRANSPARENCY AND ETHICALLY ALIGNED DESIGN

Another aspect of transparency is in the sense of traceability and verification (Winfield and Jirotka, 2017; Wortham et al., 2017). Winfield and Jirotka (2017) propose that robots should be equipped with an ethical black box, the equivalent of the black box used in aircrafts, to provide transparency about how and why a certain accident may have happened, helping to establish accountability. This transparency could help disruptive technologies gain public trust (Sciutti et al., 2018).

Ethics and Standards are interconnected, and both fit into the broader framework of Responsible Research and Innovation. There is an IEEE Global Initiative for Ethically Aligned Design for Artificial Intelligence and Autonomous Systems, with a work group dedicated to Transparency (Bryson and Winfield, 2017; Grinbaum et al., 2017). In this initiative, Transparency is defined as "the property which makes possible to discover how and why the system made a particular decision, or in the case of a robot, acted the way it did." The standard describes levels of transparency for autonomous systems for different stakeholders: users, certification agencies, accident investigators, lawyers, and general public.

European Union's new General Data Protection Regulation and the Recommendations to the Commission on Civil Law Rules on Robotics are examples of the increasing importance of ethically aligned designs. The first one creates the right to receive explanations (Goodman and Flaxman, 2016), and the second one recommends maximum transparency, predictability, and traceability (Boden et al., 2017; European Parlament, 2017).

8. DISCUSSION

Marvin Minsky used the term "suitcase word" (Minsky, 2006) to refer to words with several meanings packed into them. Transparency is a kind of suitcase-like word, so we propose a categorization of the different meanings of transparency in shared autonomy identified in the state of the art. This categorization can be found in **Table 1**.

It can be observed that algorithmic approaches to establish and improve transparency are well developed, mature, and numerous in mechanical transparency and haptic interfaces. On the other hand, algorithms to establish transparency in the sense of observability, predictability, legibility, or explainability, or for other types of interfaces like brain computer interfaces, are not so numerous and have only been recently developed. **Table 2** clusters a relevant selection of these algorithmic approaches.

Considering the challenges of transparency, several areas might be promising for future developments. The challenges of transparency in shared autonomy are different for high levels of autonomy and for low levels of autonomy.

For low levels of autonomy, the operator is doing almost everything directly, so the uncertainty and predictability are low, and transparency may be high, but there is a problem because the human cognitive workload to be aware of everything might become too high. The solutions might be:

- The use of intermediate levels of autonomy, so that the user might delegate some tasks (Miller, 2014). Trust is necessary for delegation, without trust, the user is not going to delegate, no matter how capable the robot is (Kruijff et al., 2014). Transparency helps build trust (Ososky et al., 2014).
- Improve the interfaces design to allow users to manage the information available, to obtain a high level of understanding of what is going on.
- Learn from the experience. If a robot requests human support in a difficult situation, the human actions could be stored and executed the next time the robot faces the same situation.

For high levels of autonomy, human is delegating almost everything, so the uncertainty and predictability are high, and the transparency may be low. The operator cognitive engagement and attention might become low (Endsley, 2012; Hancock, 2017), and it might cause problems detecting failures (complacency effect) (Parasuraman et al., 1993), and recovering manual control from automation failure (Lumberjack effect)(Onnasch et al., 2014; Endsley, 2017). The solutions might be:

- The use of intermediate levels of autonomy.

TABLE 2 | Main algorithmic approaches.

Transparency in shared autonomy	
3. Transparency as observability and predictability of system behavior	
Dragan et al., 2013, 2015a,b; Nikolaidis et al., 2016	Generating legible motion optimizing trajectories space
Busch et al., 2017; Buehler and Weisswange, 2018	Legibility using model-free methods
Takayama et al., 2011	Readability based on handcoded animations
Cha et al., 2017; Kim and Fong, 2017; Ganesan et al., 2018	Visual cues and light signaling
Robot-of-Human Transparency	
Bethel and Murphy, 2008; Dragan, 2017; Roncone et al., 2017; Chen et al., 2018; Gui et al., 2018	Intent recognition
Gielniak and Thomaz, 2011; Lorenz et al., 2014; Bai et al., 2015; Matthews et al., 2017; Nikolaidis et al., 2017; Wang et al., 2017; Doellinger et al., 2018; Goldhoorn et al., 2018	Mutual adaptation
Breazeal et al., 2005; Li and Zhang, 2017; Li et al., 2017; Cha et al., 2018; Gildert et al., 2018; Haji Fathaliyan et al., 2018; Lakomkin et al., 2018	Implicit communication
5. Transparency as explainability	
Caminada et al., 2014; Rosenthal et al., 2016; Wortham and Rogers, 2017; Kuhner et al., 2018; Nikolaidis et al., 2018	Route/planning/navigation verbalization
MacMahon et al., 2006; Kollar et al., 2010; Matuszek et al., 2010; Duvallet et al., 2013, 2016; Oßwald et al., 2014; Hemachandra et al., 2015; Daniele et al., 2017; Suddrey et al., 2017; Nikolaidis et al., 2018; Sinha et al., 2018	Natural language grounding
6. Mechanical transparency	
Wearables, Exoskeletons	
Robertson et al., 2007; Jarrasse et al., 2008; Jarrassé et al., 2009; van Dijk et al., 2013; Kim and Rosen, 2015; Boaventura and Buchli, 2016; Zhang et al., 2016; Awad et al., 2017; Beckerle et al., 2017; Boaventura et al., 2017; Chen et al., 2017; Fong et al., 2017; Bai et al., 2018; Fani et al., 2018	Force feedback control with feedforward loop fed with predictive information, impedance and admittance controllers, electromyography methods
Telerobotics	
Raju et al., 1989; Lawrence, 1993; Yokokohji and Yoshikawa, 1994; Baier and Schmidt, 2004; Lee and Li, 2005; Hokayem and Spong, 2006; Monfaredi et al., 2006; Ferre et al., 2007; Goethals et al., 2007; Hirche and Buss, 2007; Polushin et al., 2007; Kim et al., 2010, 2013; Yalcin and Ohnishi, 2010; Baser and Konukseven, 2012; Franken et al., 2012; Na and Vu, 2012; Slawinski et al., 2012; Aracil et al., 2013; Baser et al., 2013; Goodrich et al., 2013; Meli et al., 2014; Pacchierotti et al., 2014; Hertkorn, 2015; Farooq et al., 2016; Park et al., 2016; Sun et al., 2016; Xu et al., 2016; Gopinath et al., 2017; Lu et al., 2017	Realistic perception of the remote environment through:adaptive impedance force control, stiffness observers, position-force controllers, four channel control, impedance reflection algorithms, coupled impedance controllers, cutaneous tactile force feedback
Burget et al., 2017; Muelling et al., 2017; Zhao et al., 2017	BCI, AR

- Increase of transparency of the system's intent and reasoning, including information beyond the three levels SAT model.
- Increase robot-of-human transparency to recognize human attention reduction.
- Training to avoid the out-of-the-loop performance problem, and calibrate the trust in the system.

9. CONCLUSIONS

The current research about transparency in the shared autonomy framework has been reviewed, to provide a general and complete overview. The next ways of understanding transparency in human-robot interaction in the shared autonomy framework have been identified in the state of the art:

- Transparency as the observability of the system behavior, and as the opposite of unpredictability of the state of the system. The human understanding of what the system is doing, why, and what it will do next.
- Transparency as a method to achieve shared situation awareness and shared intent between the human and the system. The main methods to improve shared situation awareness are interface design and training.
- Robot-to-human transparency (understanding of system behavior) vs. robot-of-human transparency (understanding

of human behavior). This work has focused on the first one.

- Transparency as a property of the human-robot interface and the transparency situation awareness model. Transparent interfaces can be achieved through natural language explanations.
- Mechanical transparency used in haptics, bilateral teleoperation, and wearable robots like exoskeletons.
- Transparency as traceability and verification.

The benefits of transparency are multiple: transparency improves system performance and might reduce human errors, builds trust in the system, and transparent design principles are aligned with user-centered design, and ethically aligned design.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

REFERENCES

Anaya, F., Thangavel, P., and Yu, H. (2018). Hybrid FES–robotic gait rehabilitation technologies: a review on mechanical design, actuation, and control strategies. *Int. J. Intell. Robot. Appl.* 2, 1–28. doi: 10.1007/s41315-017-0042-6

Aracil, R., Azorin, J., Ferre, M., and Peña, C. (2013). Bilateral control by state convergence based on transparency for systems with time delay. *Robot. Auton. Syst.* 61, 86–94. doi: 10.1016/j.robot.2012.11.006

Arrichiello, F., Lillo, P. D., Vito, D. D., Antonelli, G., and Chiaverini, S. (2017). "Assistive robot operated via p300-based brain computer interface," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore), 6032–6037. doi: 10.1109/ICRA.2017.7989714

Awad, L. N., Bae, J., O'Donnell, K., De Rossi, S. M. M., Hendron, K., Sloot, L. H., et al. (2017). A soft robotic exosuit improves walking in patients after stroke. *Sci. Transl. Med.* 9:eaai9084. doi: 10.1126/scitranslmed.aai9084

Bai, H., Cai, S., Ye, N., Hsu, D., and Lee, W. S. (2015). "Intention-aware online pomdp planning for autonomous driving in a crowd," in *2015 IEEE International Conference on Robotics and Automation (ICRA)* (Seattle, WA), 454–460. doi: 10.1109/ICRA.2015.7139219

Bai, S., Sugar, T., and Virk, G. (2018). *Wearable Exoskeleton Systems: Design, Control and Applications.* Control, Robotics and Sensors Series. Institution of Engineering & Technology, UK. doi: 10.1049/PBCE108E

Baier, H., and Schmidt, G. (2004). Transparency and stability of bilateral kinesthetic teleoperation with time-delayed communication. *J. Intell. Robot. Syst.* 40, 1–22. doi: 10.1023/B:JINT.0000034338.53641.d0

Baraka, K., Rosenthal, S., and Veloso, M. (2016). "Enhancing human understanding of a mobile robot's state and actions using expressive lights," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (New York, NY), 652–657. doi: 10.1109/ROMAN.2016.7745187

Baser, O., Gurocak, H., and Konukseven, E. I. (2013). Hybrid control algorithm to improve both stable impedance range and transparency in haptic devices. *Mechatronics* 23, 121–134. doi: 10.1016/j.mechatronics.2012.11.006

Baser, O., and Konukseven, E. I. (2012). Utilization of motor current based torque feedback to improve the transparency of haptic interfaces. *Mech. Mach. Theory* 52, 78–93. doi: 10.1016/j.mechmachtheory.2012.01.012

Beckerle, P., Salvietti, G., Unal, R., Prattichizzo, D., Rossi, S., Castellini, C., et al. (2017). A human-robot interaction perspective on assistive and rehabilitation robotics. *Front. Neurorobot.* 11:24. doi: 10.3389/fnbot.2017.00024

Beer, J. M., Fisk, A. D., and Rogers, W. A. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. *J. Hum.-Robot Interact.* 3, 74–99. doi: 10.5898/JHRI.3.2.Beer

Bethel, C. L., and Murphy, R. R. (2008). Survey of non-facial/non-verbal affective expressions for appearance-constrained robots. *IEEE Trans. Syst. Man Cybern. Part C* 38, 83–92. doi: 10.1109/TSMCC.2007.905845

Bi, L., Fan, X. A., and Liu, Y. (2013). EEG-based brain-controlled mobile robots: a survey. *IEEE Trans. Hum. Mach. Syst.* 43, 161–176. doi: 10.1109/TSMCC.2012.2219046

Boaventura, T., and Buchli, J. (2016). "Acceleration-based transparency control framework for wearable robots," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Daejeon), 5683–5688. doi: 10.1109/IROS.2016.7759836

Boaventura, T., Hammer, L., and Buchli, J. (2017). "Interaction force estimation for transparency control on wearable robots using a kalman filter," in *Converging Clinical and Engineering Research on Neurorehabilitation II*, eds J. Ibáñez, J. González-Vargas, J. M. Azorín, M. Akay, and J. L. Pons (Segovia: Springer International Publishing), 489–493.

Boden, M., Bryson, J., Caldwell, D., Dautenhahn, K., Edwards, L., Kember, S., Newman, P., et al. (2017). Principles of robotics: regulating robots in the real world. *Connect. Sci.* 29, 124–129. doi: 10.1080/09540091.2016.1271400

Borenstein, J., Wagner, A. R., and he, A. (2018). Overtrust of pediatric health-care robots: a preliminary survey of parent perspectives. *IEEE Robot. Automat. Mag.* 25, 46–54. doi: 10.1109/MRA.2017.2778743

Bradshaw, J. M., Hoffman, R. R., Woods, D. D., and Johnson, M. (2013). The seven deadly myths of autonomous systems. *IEEE Intell. Syst.* 28, 54–61. doi: 10.1109/MIS.2013.70

Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., and Berlin, M. (2005). "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Edmonton, AB), 708–713. doi: 10.1109/IROS.2005.1545011

Bryson, J., and Winfield, A. (2017). Standardizing ethical design for artificial intelligence and autonomous systems. *Computer* 50, 116–119. doi: 10.1109/MC.2017.154

Buehler, M. C., and Weisswange, T. H. (2018). "Online inference of human belief for cooperative robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid), 409–415.

Burget, F., Fiederer, L. D. J., Kuhner, D., Völker, M., Aldinger, J., Schirrmeister, R. T., et al. (2017). "Acting thoughts: towards a mobile robotic service assistant for users with limited communication skills," in *2017 European Conference on Mobile Robots (ECMR)* (Paris), 1–6. doi: 10.1109/ECMR.2017.80 98658

Busch, B., Grizou, J., Lopes, M., and Stulp, F. (2017). Learning legible motion from human–robot interactions. *Int. J. Soc. Robot.* 9, 765–779. doi: 10.1007/s12369-017-0400-4

Caminada, M. W., Kutlak, R., Oren, N., and Vasconcelos, W. W. (2014). "Scrutable plan enactment via argumentation and natural language generation," in *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems* (Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems), 1625–1626.

Casalino, A., Messeri, C., Pozzi, M., Zanchettin, A. M., Rocco, P., and Prattichizzo, D. (2018). Operator awareness in humanRobot collaboration through wearable vibrotactile feedback. *IEEE Robot. Automat. Lett.* 3, 4289–4296. doi: 10.1109/LRA.2018.2865034

Cha, E., Kim, Y., Fong, T., and Mataric, M. J. (2018). A survey of nonverbal signaling methods for non-humanoid robots. *Foundat. Trends Robot.* 6, 211–323. doi: 10.1561/2300000057

Cha, E., Trehon, T., Wathieu, L., Wagner, C., Shukla, A., and Matarić, M. J. (2017). "Modlight: designing a modular light signaling tool for human-robot interaction," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore), 1654–1661. doi: 10.1109/ICRA.2017.7989195

Chang, M. L., Gutierrez, R. A., Khante, P., Short, E. S., and Thomaz, A. L. (2018). "Effects of integrated intent recognition and communication on human-robot collaboration," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid), 3381–3386.

Chen, J., Procci, K., Boyce, M., Wright, J., Garcia, A., and Barnes, M. (2014). *Situation Awareness–Based Agent Transparency*. Technical Report ARL-TR-6905, ARM US Army Research Laboratory.

Chen, M., Nikolaidis, S., Soh, H., Hsu, D., and Srinivasa, S. (2018). "Planning with trust for human-robot collaboration," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY: ACM), 307–315.

Chen, X., Zeng, Y., and Yin, Y. (2017). Improving the transparency of an exoskeleton knee joint based on the understanding of motor intent using energy kernel method of EMG. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 577–588. doi: 10.1109/TNSRE.2016.2582321

Daniele, A. F., Bansal, M., and Walter, M. R. (2017). "Navigational instruction generation as inverse reinforcement learning with neural machine translation," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, HRI '17*, (New York, NY: ACM), 109–118.

DARPA (2016). *Explainable Artificial Intelligence (xai)*. Technical Report Defense Advanced Research Projects Agency, DARPA-BAA-16-53.

Desai, M. (2012). *Modeling Trust to Improve Human-robot Interaction*. Ph.D. thesis. University of Massachusetts Lowell, Lowell, MA.

DoD (2012). *The Role of Autonomy in DoD Systems*. Technical report, Department of Defense (DoD).

Doellinger, J., Spies, M., and Burgard, W. (2018). Predicting occupancy distributions of walking humans with convolutional neural networks. *IEEE Robot. Automat. Lett.* 3, 1522–1528. doi: 10.1109/LRA.2018.2800780

Dragan, A. D. (2017). "Robot planning with mathematical models of human state and action," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, Workshop in User Centered Design* (Madrid). Available online at: http://arxiv.org/abs/1705.04226

Dragan, A. D., Bauman, S., Forlizzi, J., and Srinivasa, S. S. (2015a). "Effects of robot motion on human-robot collaboration," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY: ACM), 51–58.

Dragan, A. D., Lee, K. C., and Srinivasa, S. S. (2013). "Legibility and predictability of robot motion," in *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction, HRI '13* (Piscataway, NJ: IEEE Press), 301–308.

Dragan, A. D., Muelling, K., Bagnell, J. A., and Srinivasa, S. S. (2015b). "Movement primitives via optimization," in *2015 IEEE International Conference on Robotics and Automation (ICRA)* (Seattle, WA), 2339–2346. doi: 10.1109/ICRA.2015.7139510

Duvallet, F., Kollar, T., and Stentz, A. (2013). "Imitation learning for natural language direction following through unknown environments," in *2013 IEEE International Conference on Robotics and Automation* (Karlsruhe), 1047–1053. doi: 10.1109/ICRA.2013.6630702

Duvallet, F., Walter, M. R., Howard, T., Hemachandra, S., Oh, J., Teller, S., et al. (2016). *Inferring Maps and Behaviors from Natural Language Instructions*. Cham: Springer International Publishing.

Endsley, M. (2012). *Designing for Situation Awareness: An Approach to User-Centered Design, 2nd Edition*. Boca Raton, FL: CRC Press.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors* 37, 32–64.

Endsley, M. R. (2017). From here to autonomy: lessons learned from human-automation research. *Human Factors* 59, 5–27. doi: 10.1177/0018720816681350

Endsley, M. R. (2018). Level of automation forms a key aspect of autonomy design. *J. Cogn. Eng. Decis. Making* 12, 29–34. doi: 10.1177/1555343417723432

Endsley, M. R., and Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics* 42, 462–492.

European Parlament (2017). *Report With Recommendations to the Commission on Civil Law Rules on Robotics*. Technical Report 2015/2103(INL), Committee on Legal Affairs.

Ezeh, C., Trautman, P., Devigne, L., Bureau, V., Babel, M., and Carlson, T. (2017). "Probabilistic vs linear blending approaches to shared control for wheelchair driving," in *2017 International Conference on Rehabilitation Robotics (ICORR)* (London), 835–840. doi: 10.1109/ICORR.2017.8009352

Fani, S., Ciotti, S., Catalano, M. G., Grioli, G., Tognetti, A., Valenza, G., et al. (2018). Simplifying telerobotics: wearability and teleimpedance improves human-robot interactions in teleoperation. *IEEE Robot. Automat. Mag.* 25, 77–88. doi: 10.1109/MRA.2017.2741579

Farooq, U., Gu, J., El-Hawary, M., Asad, M. U., and Rafiq, F. (2016). "Transparent fuzzy bilateral control of a nonlinear teleoperation system through state convergence," in *2016 International Conference on Emerging Technologies (ICET)* (Islamabad), 1–6. doi: 10.1109/ICET.2016.7813242

Ferre, M., Buss, M., Aracil, R., Melchiorri, C., and Balaguer, C. (2007). *Advances in Telerobotics*. Springer Berlin; Heidelberg: Springer Tracts in Advanced Robotics.

Fischer, K., Weigelin, H. M., and Bodenhagen, L. (2018). Increasing trust in human-robot medical interactions: effects of transparency and adaptability. *Paladyn* 9, 95–109. doi: 10.1515/pjbr-2018-0007

Fong, J., Crocher, V., Tan, Y., Oetomo, D., and Mareels, I. (2017). "EMU: A transparent 3D robotic manipulandum for upper-limb rehabilitation," in *2017 International Conference on Rehabilitation Robotics (ICORR)* (London), 771–776. doi: 10.1109/ICORR.2017.8009341

Franken, M., Misra, S., and Stramigioli, S. (2012). Improved transparency in energy-based bilateral telemanipulation. *Mechatronics* 22, 45–54. doi: 10.1016/j.mechatronics.2011.11.004

Ganesan, R. K., Rathore, Y. K., Ross, H. M., and Amor, H. B. (2018). Better teaming through visual cues: how projecting imagery in a workspace can improve human-robot collaboration. *IEEE Robot. Automat. Mag.* 25, 59–71. doi: 10.1109/MRA.2018.2815655

Gielniak, M. J., and Thomaz, A. L. (2011). "Generating anticipation in robot motion," in *2011 RO-MAN* (Atlanta, GA), 449–454. doi: 10.1109/ROMAN.2011.6005255

Gildert, N., Millard, A. G., Pomfret, A., and Timmis, J. (2018). The need for combining implicit and explicit communication in cooperative robotic systems. *Front. Robot.* 5:65. doi: 10.3389/frobt.2018.00065

Goethals, P., Gersem, G. D., Sette, M., and Reynaerts, D. (2007). "Accurate haptic teleoperation on soft tissues through slave friction compensation by impedance reflection," in *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07)* (Tsukaba), 458–463. doi: 10.1109/WHC.2007.17

Goldhoorn, A., Garrell, A., Renéand A., and Sanfeliu, A. (2018). Searching and tracking people with cooperative mobile robots. *Auton. Robots* 42, 739–759. doi: 10.1007/s10514-017-9681-6

Goodman, B., and Flaxman, S. (2016). European Union Regulations on algorithmic decision-making and a "Right to Explanation". *AI Magazine* 38, 50–57. doi: 10.1609/aimag.v38i3.2741

Goodrich, M., W. Crandall, J., and Barakova, E. (2013). Teleoperation and beyond for assistive humanoid robots. *Rev. Hum. Factors Ergon.* 9, 175–226. doi: 10.1177/1557234X13502463

Goodrich, M. A., and Schultz, A. C. (2007). Human-robot interaction: a survey. *Found. Trends Hum. Comput. Interact.* 1, 203–275. doi: 10.1561/1100000005

Gopinath, D., Jain, S., and Argall, B. D. (2017). Human-in-the-loop optimization of shared autonomy in assistive robotics. *IEEE Robot. Automat. Lett.* 2, 247–254. doi: 10.1109/LRA.2016.2593928

Gransche, B., Shala, E., Hubig, C., Alpsancar, S., and Harrach, S., (2014). *Wande von Autonomie und Kontrolle durch neue Mensch-Technik-Interaktionen:Grundsatzfragen autonomieorienter*. Stuttgart: Mensch-Technik-Verhältnisse.

Grinbaum, A., Chatila, R., Devillers, L., Ganascia, J. G., Tessier, C., and Dauchet, M. (2017). Ethics in robotics research: Cerna mission and context. *IEEE Robot. Automat. Mag.* 24, 139–145. doi: 10.1109/MRA.2016.2611586

Gui, L.-Y., Zhang, K., Wang, Y.-X., Liang, X., Moura, J. M. F., and Veloso, M. (2018). "Teaching robots to predict human motion," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid), 562–567.

Haji Fathaliyan, A., Wang, X., and Santos, V. J. (2018). Exploiting three-dimensional gaze tracking for action recognition during bimanual

manipulation to enhance humanRobot collaboration. *Front. Robot. AI* 5:25. doi: 10.3389/frobt.2018.00025

Hancock, P. A. (2017). On the nature of vigilance. *Hum. Factors* 59, 35–43. doi: 10.1177/0018720816655240

Hayes, B., and Shah, J. A. (2017). "Improving robot controller transparency through autonomous policy explanation," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY: ACM), 303–312.

Hellström, T., and Bensch, S. (2018). Understandable robots. *Paladyn J. Behav. Robot.* 9, 110–123. doi: 10.1515/pjbr-2018-0009

Hemachandra, S., Duvallet, F., Howard, T. M., Roy, N., Stentz, A., and Walter, M. R. (2015). "Learning models for following natural language directions in unknown environments," in *2015 IEEE International Conference on Robotics and Automation (ICRA)* (Seattle, WA), 5608–5615. doi: 10.1109/ICRA.2015.7139984

Hertkorn, K. (2015). *Shared Grasping: a Combination of Telepresence and Grasp Planning*. Ph.D. thesis, Karlsruher Institut für Technologie (KIT).

Hirche, S., and Buss, M. (2007). "Human perceived transparency with time delay," in *Advances in Telerobotics*, eds M. Ferre, M. Buss, R. Aracil, C. Melchiorri, and C. Balaguer (Berlin; Heidelberg: Springer), 191–209. doi: 10.1007/978-3-540-71364-7_13

Hokayem, P. F., and Spong, M. W. (2006). Bilateral teleoperation: An historical survey. *Automatica* 42, 2035–2057. doi: 10.1016/j.automatica.2006.06.027

Iden, J. (2017). "Belief, judgment, transparency, trust: reasoning about potential pitfalls in interacting with artificial autonomous entities," in *Robotics: Science and Systems XIII, RSS 2017*, eds N. Amato, S. Srinivasa, and N. Ayanian (Cambridge, MA).

Jarrasse, N., Paik, J., Pasqui, V., and Morel, G. (2008). "How can human motion prediction increase transparency?" in *2008 IEEE International Conference on Robotics and Automation* (Pasadena, CA), 2134–2139. doi: 10.1109/ROBOT.2008.4543522

Jarrassé, N., Paik, J., Pasqui, V., and Morel, G. (2009). *Experimental Evaluation of Several Strategies for Human Motion Based Transparency Control*, pages 557–565. Berlin; Heidelberg: Springer Berlin Heidelberg.

Javdani, S., Admoni, H., Pellegrinelli, S., Srinivasa, S. S., and Bagnell, J. A. (2018). Shared autonomy via hindsight optimization for teleoperation and teaming. *Int. J. Robot. Res.* 37, 717–742. doi: 10.1177/0278364918776060

Jones, D. G., and Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviat. Space Environ. Med.* 67, 507–512.

Kaber, D. B. (2017). Issues in human-automation interaction modeling: presumptive aspects of frameworks of types and levels of automation. *J. Cogn. Eng. Decis. Mak.* 12:155534341773720. doi: 10.1177/1555343417737203

Kim, H., and Rosen, J. (2015). Predicting redundancy of a 7 dof upper limb exoskeleton toward improved transparency between human and robot. *J. Intell. Robot. Syst.* 80, 99–119. doi: 10.1007/s10846-015-0212-4

Kim, J., Chang, P. H., and Park, H. (2013). Two-channel transparency-optimized control architectures in bilateral teleoperation with time delay. *IEEE Trans. Control Syst. Technol.* 21, 40–51. doi: 10.1109/TCST.2011.2172945

Kim, J., Park, H.-S., and Chang, P. H. (2010). Simple and robust attainment of transparency based on two-channel control architectures using time-delay control. *J. Intell. Robot. Syst.* 58, 309–337. doi: 10.1007/s10846-009-9376-0

Kim, T., and Hinds, P. (2006). "Who should i blame? Effects of autonomy and transparency on attributions in human-robot interaction," in *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication* (Hatfield), 80–85. doi: 10.1109/ROMAN.2006.314398

Kim, Y., and Fong, T. (2017). "Signaling robot state with light attributes," in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY: ACM), 163–164.

Kollar, T., Tellex, S., Roy, D., and Roy, N. (2010). "Toward understanding natural language directions," in *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction* (Piscataway, NJ: IEEE Press), 259–266.

Kruijff, G. J. M., Janíček, M., Keshavdas, S., Larochelle, B., Zender, H., Smets, N. J. J. M., et al. (2014). *Experience in System Design for Human-Robot Teaming in Urban Search and Rescue*. Berlin; Heidelberg: Springer Berlin Heidelberg.

Kuhner, D., Aldinger, J., Burget, F., Göbelbecker, M., Burgard, W., and Nebel, B. (2018). "Closed-loop robot task planning based on referring expressions," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid), 876–881.

Kuipers, B. (2018). How can we trust a robot? *Commun. ACM* 61, 86–95. doi: 10.1145/3173087

Lakhmani, S., Abich, J., Barber, D., and Chen, J. (2016). *A Proposed Approach for Determining the Influence of Multimodal Robot-of-Human Transparency Information on Human-Agent Teams*. Cham: Springer International Publishing.

Lakomkin, E., Zamani, M. A., Weber, C., Magg, S., and Wermter, S. (2018). "Emorl: continuous acoustic emotion classification using deep reinforcement learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (Brisbane, QLD), 1–6. doi: 10.1109/ICRA.2018.8461058

Lawrence, D. A. (1993). Stability and transparency in bilateral teleoperation. *IEEE Trans. Robot. Automat.* 9, 624–637.

Lee, D., and Li, P. Y. (2005). Passive bilateral control and tool dynamics rendering for nonlinear mechanical teleoperators. *IEEE Transact. Robot.* 21, 936–951. doi: 10.1109/TRO.2005.852259

Lee, J. D., and See, K. A. (2004). Trust in automation: designing for appropriate reliance. *Hum. Factors* 46, 50–80. doi: 10.1518/hfes.46.1.50.30392

Lewis, M., Sycara, K., and Walker, P. (2018). *The Role of Trust in Human-Robot Interaction*. Cham: Springer International Publishing. doi: 10.1007/978-3-319-64816-3_8

Li, S., and Zhang, X. (2017). Implicit intention communication in human-robot interaction through visual behavior studies. *IEEE Trans. Hum. Mach. Syst.* 47, 437–448. doi: 10.1109/THMS.2017.2647882

Li, S., Zhang, X., and Webb, J. D. (2017). 3-D-gaze-based robotic grasping through mimicking human visuomotor function for people with motion impairments. *IEEE Trans. Biomed. Eng.* 64, 2824–2835.

Lorenz, T. (2015). *Emergent Coordination Between Humans and Robots*. Ph.D. thesis, Ludwig-Maximilians-Universität München.

Lorenz, T., Vlaskamp, B. N. S., Kasparbauer, A.-M., Mörtl, A., and Hirche, S. (2014). Dyadic movement synchronization while performing incongruent trajectories requires mutual adaptation. *Front. Hum. Neurosci.* 8:461. doi: 10.3389/fnhum.2014.00461

Lu, Z., Huang, P., Dai, P., Liu, Z., and Meng, Z. (2017). Enhanced transparency dual-user shared control teleoperation architecture with multiple adaptive dominance factors. *Int. J. Control Automat. Syst.* 15, 2301–2312. doi: 10.1007/s12555-016-0467-y

Lyons, J. (2013). "Being transparent about transparency: a model for human-robot interaction," in *AAAI Spring Symposium* (Palo Alto, CA), 48–53.

Lyons, J. B., and Havig, P. R. (2014). *Transparency in a Human-Machine Context: Approaches for Fostering Shared Awareness/Intent*. Cham: Springer International Publishing, 181–190.

MacMahon, M., Stankiewicz, B., and Kuipers, B. (2006). "Walk the talk: connecting language, knowledge, and action in route instructions," in *Proceedings of the 21st National Conference on Artificial Intelligence - Volume 2* (Boston, MA: AAAI Press), 1475–1482.

Matthews, M., Chowdhary, G., and Kieson, E. (2017). "Intent communication between autonomous vehicles and pedestrians," in *Proceedings of the Robotics: Science and Systems XI, RSS 2015*, eds L. E. Kavraki, D. Hsu, and J. Buchli (Rome). Available online at: http://arxiv.org/abs/1708.07123

Matuszek, C., Fox, D., and Koscher, K. (2010). "Following directions using statistical machine translation," in *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction* (Piscataway, NJ: IEEE Press), 251–258.

Meli, L., Pacchierotti, C., and Prattichizzo, D. (2014). Sensory subtraction in robot-assisted surgery: fingertip skin deformation feedback to ensure safety and improve transparency in bimanual haptic interaction. *IEEE Trans. Biomed. Eng.* 61, 1318–1327. doi: 10.1109/TBME.2014.2303052

Miller, C. A. (2014). *Delegation and Transparency: Coordinating Interactions So Information Exchange Is No Surprise*. Cham: Springer International Publishing, 191–202.

Miller, C. A. (2017). The risks of discretization: what is lost in (even good) levels-of-automation schemes. *J. Cogn. Eng. Decis. Mak.* 12, 74–76. doi: 10.1177/1555343417726254

Miller, C. A. (2018). "Displaced interactions in human-automation relationships: Transparency over time," in *Engineering Psychology and Cognitive Ergonomics*, ed D. Harris (Cham. Springer International Publishing), 191–203.

Minsky, M. (2006). *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York, NY: Simon & Schuster, Inc.

Monfaredi, R., Razi, K., Ghydari, S. S., and Rezaei, S. M. (2006). "Achieving high transparency in bilateral teleoperation using stiffness observer for passivity control," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Beijing), 1686–1691. doi: 10.1109/IROS.2006.282125

Mueller, E. (2016). *Transparent Computers: Designing Understandable Intelligent Systems*. Scotts Valley, CA: CreateSpace Independent Publishing Platform.

Muelling, K., Venkatraman, A., Valois, J.-S., Downey, J. E., Weiss, J., Javdani, S., et al. (2017). Autonomy infused teleoperation with application to brain computer interface controlled manipulation. *Auton. Robots* 41, 1401–1422. doi: 10.1007/s10514-017-9622-4

Murphy, R. R. (2014). *Disaster Robotics*. Cambridge, MA: The MIT Press.

Na, U. J., and Vu, M. H. (2012). "Adaptive impedance control of a haptic teleoperation system for improved transparency," in *2012 IEEE International Workshop on Haptic Audio Visual Environments and Games (HAVE 2012) Proceedings* (Munich), 38–43. doi: 10.1109/HAVE.2012.6374442

Nikolaidis, S., Dragan, A., and Srinivasa, S. (2016). "Viewpoint-based legibility optimization," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Christchurch), 271–278. doi: 10.1109/HRI.2016.7451762

Nikolaidis, S., Kwon, M., Forlizzi, J., and Srinivasa, S. (2018). Planning with verbal communication for human-robot collaboration. *ACM Trans. Human Robot Interact.* 7, 22:1–22:21.

Nikolaidis, S., Lasota, P., Ramakrishnan, R., and Shah, J. (2015). Improved human-robot team performance through cross-training, an approach inspired by human team training practices. *Int. J. Robot. Res.* 34, 1711–1730. doi: 10.1177/0278364915609673

Nikolaidis, S., Zhu, Y. X., Hsu, D., and Srinivasa, S. (2017). "Human-robot mutual adaptation in shared autonomy," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY: ACM), 294–302.

Oßwald, S., Kretzschmar, H., Burgard, W., and Stachniss, C. (2014). "Learning to give route directions from human demonstrations," in *2014 IEEE International Conference on Robotics and Automation (ICRA)* (Hong Kong), 3303–3308. doi: 10.1109/ICRA.2014.6907334

Okamura, A. M. (2018). Haptic dimensions of human-robot interaction. *ACM Trans. Hum. Robot Interact.* 7, 6:1–6:3.

Onnasch, L., Wickens, C. D., Li, H., and Manzey, D. (2014). Human performance consequences of stages and levels of automation: an integrated meta-analysis. *Hum. Factors* 56, 476–488. doi: 10.1177/0018720813501549

Ososky, S., Sanders, T., Jentsch, F., Hancock, P., and Chen, J. (2014). "Determinants of system transparency and its influence on trust in and reliance on unmanned robotic systems," in *Conference SPIE Defense and Security* (Baltimore, MA), 9084–9096. doi: 10.1117/12.2050622

Oviatt, S., Schuller, B., Cohen, P. R., Sonntag, D., Potamianos, G., and Krüger, A. (eds.). (2017). *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1.* New York, NY: Association for Computing Machinery and Morgan.

Pacchierotti, C., Tirmizi, A., and Prattichizzo, D. (2014). Improving transparency in teleoperation by means of cutaneous tactile force feedback. *ACM Trans. Appl. Percept.* 11:4. doi: 10.1145/2604969

Parasuraman, R., Molloy, R., and Singh, I. (1993). Performance consequences of automation induced complacency. *Int. J. Aviat. Psychol.* 3, 1–23.

Parasuraman, R., and Riley, V. (1997). Humans and automation: use, misuse, disuse, abuse. *Hum. Factors* 39, 230–253.

Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Trans. Syst. Man Cybernet.* 30, 286–297. doi: 10.1109/3468.844354

Park, S., Uddin, R., and Ryu, J. (2016). Stiffness-reflecting energy-bounding approach for improving transparency of delayed haptic interaction systems. *Int. J. Control Automat. Syst.* 14, 835–844. doi: 10.1007/s12555-014-0109-9

Perera, V., Selvaraj, S. P., Rosenthal, S., and Veloso, M. (2016). "Dynamic generation and refinement of robot verbalization," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (New York, NY), 212–218. doi: 10.1109/ROMAN.2016.7745133

Perzanowski, D., Schultz, A. C., Adams, W., Marsh, E., and Bugajska, M. (2001). Building a multimodal human-robot interface. *IEEE Intell. Syst.* 16, 16–21. doi: 10.1109/MIS.2001.1183338

Polushin, I. G., Liu, P. X., and Lung, C. (2007). A force-reflection algorithm for improved transparency in bilateral teleoperation with communication delay. *IEEE/ASME Trans. Mechatron.* 12, 361–374. doi: 10.1109/TMECH.2007.897285

Raju, G. J., Verghese, G. C., and Sheridan, T. B. (1989). "Design issues in 2-port network models of bilateral remote manipulation," in *Proceedings, 1989 International Conference on Robotics and Automation* (Scottsdale, AZ), 1316–1321. doi: 10.1109/ROBOT.1989.100162

Robertson, J., Jarrassé, N., Pasqui, V., and Roby-Brami, A. (2007). De l'utilisation des robots pour la rééducation: intérêt et perspectives. *La Lettre de méDecine Phys. de Réadaptation*, 23, 139–147. doi: 10.1007/s11659-007-0070-y

Roncone, A., Mangin, O., and Scassellati, B. (2017). "Transparent role assignment and task allocation in human robot collaboration," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore), 1014–1021. doi: 10.1109/ICRA.2017.7989122

Rosenthal, S., Selvaraj, S. P., and Veloso, M. (2016). "Verbalization: narration of autonomous robot experience," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence* (New York, NY: AAAI Press), 862–868.

Rupp, R., Kleih, S. C., Leeb, R., del R. Millan, J., Kübler, A., and Müller-Putz, G. R. (2014). *Brain–Computer Interfaces and Assistive Technology*. Dordrecht: Springer Netherlands.

Sanders, T. L., Wixon, T., Schafer, K. E., Chen, J. Y. C., and Hancock, P. A. (2014). "The influence of modality and transparency on trust in human-robot interaction," in *2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)* (San Antonio, TX), 156–159. doi: 10.1109/CogSIMA.2014.6816556

Schilling, M., Kopp, S., Wachsmuth, S., Wrede, B., Ritter, H., Brox, T., et al. (2016). "Towards a multidimensional perspective on shared autonomy," in *Proceedings of the AAAI Fall Symposium Series 2016* (Stanford, CA).

Sciutti, A., Mara, M., Tagliasco, V., and Sandini, G. (2018). Humanizing human-robot interaction: on the importance of mutual understanding. *IEEE Technol. Soc. Mag.* 37, 22–29. doi: 10.1109/MTS.2018.2795095

Sheh, R. K. (2017). "Why did you do that? Explainable intelligent robots," in *AAAI-17 Whorkshop on Human Aware Artificial Intelligence* (San Francisco, CA), 628–634.

Sheridan, T., and Verplank, W. (1978). *Human and Computer Control of Undersea Teleoperators*. Technical report, Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT.

Sinha, A, Akilesh, B., Sarkar, M., and Krishnamurthy, B. (2018). "Attention based natural language grounding by navigating virtual environment," in *Applications of Computer Vision, WACV 19, Hawaii*. Available online at: http://arxiv.org/abs/1804.08454

Slawinski, E., Mut, V. A., Fiorini, P., and Salinas, L. R. (2012). Quantitative absolute transparency for bilateral teleoperation of mobile robots. *IEEE Trans. Syst. Man Cybernet. Part A*, 42, 430–442. doi: 10.1109/TSMCA.2011.2159588

Suddrey, G., Lehnert, C., Eich, M., Maire, F., and Roberts, J. (2017). Teaching robots generalizable hierarchical tasks through natural language instruction. *IEEE Robot. Automat. Lett.* 2, 201–208. doi: 10.1109/LRA.2016.2588584

Sun, D., Naghdy, F., and Du, H. (2016). A novel approach for stability and transparency control of nonlinear bilateral teleoperation system with time delays. *Control Eng. Practice* 47, 15–27. doi: 10.1016/j.conengprac.2015.11.003

Takayama, L., Dooley, D., and Ju, W. (2011). "Expressing thought: improving robot readability with animation principles," in *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Lausanne), 69–76. doi: 10.1145/1957656.1957674

Theodorou, A., Wortham, R. H., and Bryson, J. J. (2016). "Why is my robot behaving like that? Designing transparency for real time inspection of autonomous robots," in *AISB Workshop on Principles of Robotics* (Sheffield).

Theodorou, A., Wortham, R. H., and Bryson, J. J. (2017). Designing and implementing transparency for real time inspection of autonomous robots. *Connect. Sci.* 29, 230–241. doi: 10.1080/09540091.2017.1310182

Tsiourti, C., and Weiss, A. (2014). "Multimodal affective behaviour expression: Can it transfer intentions?," in *Conference on Human-Robot Interaction (HRI2017)* (Vienna).

van Dijk, W., van der Kooij, H., Koopman, B., van Asseldonk, E. H. F., and van der Kooij, H. (2013). "Improving the transparency of a rehabilitation

robot by exploiting the cyclic behaviour of walking," in *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)* (Bellevue, WA), 1–8. doi: 10.1109/ICORR.2013.6650393

Villani, V., Sabattini, L., Czerniak, J. N., Mertens, A., and Fantuzzi, C. (2018). MATE robots simplifying my work: the benefits and socioethical implications. *IEEE Robot. Automat. Mag.* 25, 37–45. doi: 10.1109/MRA.2017.2781308

Walker, M., Hedayati, H., Lee, J., and Szafir, D. (2018). "Communicating robot motion intent with augmented reality," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid), 316–324.

Wang, Z., Boularias, A., Mülling, K., Schölkopf, B., and Peters, J. (2017). Anticipatory action selection for human-robot table tennis. *Artif. Intell.* 247, 399–414. doi: 10.1016/j.artint.2014.11.007

Westlund, J. M. K., and Breazeal, C. (2016). "Transparency, teleoperation, and children's understanding of social robots," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Christchurch), 625–626.

Winfield, A. F. T., and Jirotka, M. (2017). "The case for an ethical black box," in *Towards Autonomous Robotic Systems: 18th Annual Conference, TAROS 2017*, eds Y. Gao, S. Fallah, Y. Jin, and C. Lekakou (Guildford: Springer International Publishing), 262–273. doi: 10.1007/978-3-319-64107-2_21

Wortham, R. H., and Rogers, V. (2017). "The muttering robot: improving robot transparency though vocalisation of reactive plan execution," in *26th IEEE International Symposium on Robot and Human Interactive Communication (Ro-Man) Workshop on Agent Transparency for Human-Autonomy Teaming Effectiveness* (Lisbon).

Wortham, R. H., Theodorou, A., and Bryson, J. J. (2016). "What does the robot think? Transparency as a fundamental design requirement for intelligent systems," in *Proceedings of the IJCAI Workshop on Ethics for Artificial Intelligence* (New York, NY).

Wortham, R. H., Theodorou, A., and Bryson, J. J. (2017). "Robot transparency: improving understanding of intelligent behaviour for designers and users," in *Towards Autonomous Robotic Systems: 18th Annual Conference, TAROS 2017*, eds Y. Gao, S. Fallah, Y. Jin, and C. Lekakou (Guildford: Springer International Publishing), 274–289. doi: 10.1007/978-3-319-64107-2_22

Wright, J. L., Chen, J. Y., Barnes, M. J., and Hancock, P. A. (2017). *Agent Reasoning Transparency: The Influence of Information Level on Automation-Induced Complacency.* Technical Report ARL-TR-8044, ARM US Army Research Laboratory.

Xu, X., Cizmeci, B., Schuwerk, C., and Steinbach, E. (2016). Model-mediated teleoperation: toward stable and transparent teleoperation systems. *IEEE Access* 4, 425–449. doi: 10.1109/ACCESS.2016.2517926

Yalcin, B., and Ohnishi, K. (2010). Stable and transparent time-delayed teleoperation by direct acceleration waves. *IEEE Trans. Indus. Electron.* 57, 3228–3238. doi: 10.1109/TIE.2009.2038330

Yang, X. J., Unhelkar, V. V., Li, K., and Shah, J. A. (2017). "Evaluating effects of user experience and system transparency on trust in automation," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY: ACM), 408–416.

Yokokohji, Y., and Yoshikawa, T. (1994). Bilateral control of master-slave manipulators for ideal kinesthetic coupling-formulation and experiment. *IEEE Trans. Robot. Automat.* 10, 605–620.

Zhang, W., White, M., Zahabi, M., Winslow, A. T., Zhang, F., Huang, H., et al. (2016). "Cognitive workload in conventional direct control vs. pattern recognition control of an upper-limb prosthesis," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (Budapest), 2335–2340. doi: 10.1109/SMC.2016.7844587

Zhao, Z., Huang, P., Lu, Z., and Liu, Z. (2017). Augmented reality for enhancing tele-robotic system with force feedback. *Robot. Auton. Syst.* 96, 93–101. doi: 10.1016/j.robot.2017.05.017

Zhu, Q., Perera, V., Wächter, M., Asfour, T., and Veloso, M. (2017). "Autonomous narration of humanoid robot kitchen task experience," in *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)* (Birmingham), 390–397. doi: 10.1109/HUMANOIDS.2017.82 46903

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.