Handbook of Human Factors and Ergonomics, Fifth Edition. Gavriel Salvendy and Waldemar Karwowski.

CHAPTER 44

HUMAN–ROBOT INTERACTION

Jessie Y.C. Chen and Michael J. Barnes
U.S. Army Research Laboratory
Adelphi, Maryland

1 INTRODUCTION

The relationship of humans to robots has evolved from using tools, to controlling platforms, to supervising automated systems, to collaborating with autonomous agents (as either embodied robots or disembodied algorithms) (Chen & Barnes 2014; Chen, Barnes, & Harper-Sciarini, 2011; Chen, Haas, & Barnes, 2007; Goodrich & Schultz, 2007; Murphy, 2014; Sheridan, 2016; Steinfeld et al., 2006; Yanco & Drury 2004). One notable example of a large initiative on human-robot interaction (HRI) and collaboration is the US National Science Foundation's National Robotics Initiative (NRI 2.0: Ubiquitous Collaborative Robots), which promotes research on human collaboration with robots ("co-robots") to accomplish tasks in diverse settings such as homes, schools, healthcare, manufacturing, agriculture, energy, mining, military, to name just a few (National Science Foundation, 2017). The human factor's evolution is driven by the increasing intelligence and capability evinced in autonomous vehicles, military robots, machine learning, and human-agent communications (Chen et al., 2018; Gombolay, Huang, & Shah, 2015; Yanco & Drury, 2004). Shared control among humans and intelligent systems creates new issues in human factors, such as mutual trust as well as issues related to ethical and legal responsibility, which by definition, reside with the human (Cummings & Britton, 2019).

Early research on the human factors of automation identified automation principles that help form the basis of our current understanding of human interaction with autonomous systems (Parasuraman, Sheridan, & Wickens, 2000; Sheridan, 1992). Over the years, interface design guidance has been put forward to enhance calibrated trust and thus reduce *misuse* and *disuse* of the system by supporting transparency in terms of system purpose, process, and performance (Chen et al., 2018; Lee & See, 2004; Parasuraman & Riley, 1997). However, during real-world operations, particularly those involving high-stake missions, such as military and disaster relief, human operators often have to deal with multiple tasks besides working with the robots. Issues such as operator complacency and being *out-of-the-loop* (Endsley & Kiris, 1995; Parasuraman & Manzey, 2010) have been studied extensively in the human factors community, and ways to deal with these human attention issues have provided useful guidance to designing effective human-robot interfaces. For example, *adaptive* systems are designed to keep operators *on-the-loop* by requiring humans to be engaged in a task except during high-workload mission segments and then returning the task to manual mode when the workload is deemed manageable (Chen & Barnes, 2014; Parasuraman, Barnes, & Cosenzo, 2007). *Adaptable* systems, by contrast, keep operators actively engaged and allow them to decide when and how automation should be triggered (Miller & Parasuraman, 2007). An even more flexible and fluid collaborative scheme, mixed-initiative control has also been developed to support collaborative processes in which each partner contributes to an ongoing task based on their strengths for that task (Barnes, Chen, & Jentsch, 2015; Goodrich & Schultz, 2007; Jiang & Arkin, 2015).

The distinction between autonomous and automated systems is not always clear-cut; automated systems are more task-specific with deterministic outcomes (process control); autonomous systems usually consist of a number of inter-related automated components that are designed to make decisions whose outcomes are uncertain (e.g., collision avoidance) (Bhaskara, Skinner, & Loft, 2020). The differences have important HRI implications: *automated* systems designers focus on specific types/levels of automation and its efficacy for a particular task but also on the relationship of automation to an operator's cognitive processing requirements (Parasuraman et al., 2000; Wright, Chen, & Barnes, 2018); *supervisory control* is concerned mainly with process and performance and ensuring that tasks are being performed as specified and that the underlying processing are functioning correctly for the current situation (Chen, Barnes, & Harper-Sciarini, 2011; Sheridan, 1992; Parasuraman & Manzey, 2010); with *autonomous* systems, operators are more concerned with oversight and focus on future trajectories and uncertainty (Chen & Barnes, 2014).

As systems become increasingly sophisticated and powerful and, especially in time-constrained situations, start to make decisions autonomously, effective human-system collaboration has become more important than ever. Just as senior and junior team members must learn when to take action, collaborate, or defer to their partner depending on the unfolding dynamics; agent architectures will require two-way understanding among agents and humans to engender mutual trust that is appropriately calibrated for specific situations (Chen et al., 2018; Cooke, 2015; Lee & See, 2004; Parasuraman & Riley, 1997). Recent improvements in artificial intelligence and machine learning have improved the ability of agents to be aware of their environment, to improve their problem-solving capabilities, and to be able to communicate with their human partners (Barnes, Lakhmani, et al., 2019; Pynadath, Barnes, et al., 2018; Wang et al., 2018). Issues such as shared mental models, natural language interaction, uncertainty, and predictions of future states will become essential considerations as autonomous systems become more common (Chen & Barnes, 2014; Holder, 2018; Pollard et al., 2018; Rous, Cannon-Bowers, & Salas, 1992; Wang et al., 2018; Yena et al., 2006).

The chapter discusses the human element of the evolving man-machine symbiosis including both traditional and emerging human factors issues related to designing systems that emulate human teams operating in complex decision spaces (Chen et al., 2018; Chen & Barnes, 2014; Lyons et al., 2018). The chapter will review common HRI techniques for robots with various degrees of autonomy, including conventional, multimodal, and advanced techniques such as augmented-reality-based systems. Detailed discussion of human-robot communications will review the state-of-the-art and promising techniques currently under development. Particularly, communications-related issues such as agent transparency will be examined in detail. Finally, a thorough examination of human performance issues (situation awareness, trust, workload, training, and individual differences) will be provided.

2 HUMAN–ROBOT INTERACTION APPLICATIONS AND TECHNIQUES

The first section of this chapter reviews the state-of-the-art of HRI applications and techniques. Section 2.1 discusses manual control—direct (or near direct) control; within line of sight or short-range teleoperation; and long-range teleoperation. Common techniques for humans to manually interact with robots are discussed: joystick and game controller; multimodal interfaces (voice, haptic, and gestural); and virtual, augmented, and mixed reality (VAMR). Section 2.2 discusses human interaction with a single (semi-)autonomous robot, including both humanoid and nonhumanoid platforms. Common HRI techniques include waypoint operations, learning from human demonstration, and natural language interaction. Finally, Section 2.3 examines human interaction with multiple robots, including visual assistant and marsupial systems, multiagent systems, and swarms. Common HRI techniques include supervisory control paradigms and automated planners.

2.1 Single Robot: Manual Control and Teleoperation

Sheridan (1992) provides a widely-adopted definition of teleoperation (“The term *teleoperation* most commonly refers to direct and continuous control of a [robot]” p. 4), although the distance between the human operator and the robot can vary greatly and the sizes of the systems range from nanoscale objects to multi-ton vehicles. Manually controlled and teleoperated robots have been used in a wide variety of situations, ranging from exoskeleton and rehab, robotic surgeries, agriculture, entertainment, search and rescue activities related to disasters, military missions (e.g., aerial surveillance and reconnaissance, hazardous materials detection and disposal), undersea operations, to space applications. While many robotic systems have (semi-)autonomous capabilities, manual control and teleoperations are essential elements of these systems for the foreseeable future (either by design or when semiautonomous robots need human intervention). For example, manual control is necessary when semiautonomous systems encounter particularly difficult terrain, including natural or manmade obstacles. Even for systems that purport to be fully autonomous, manual control will be a default mode for robotic systems that have significant military roles or safety issues. Typical operator control units include panels displaying: (1) sensor view and/or data transmitted from the robot; (2) plans and commands issued to the robot; (3) status of the robot (e.g., health information about the battery and sensors, pose of the robot); (4) status of the tasks; and (5) map displays to maintain the operator’s situation awareness and to facilitate navigation. These systems often require the human controller to operate in diverse environments, with diverse systems under difficult and, in some cases, stressful situations.

2.1.1 Applications

Direct (or Near-Direct) Control Robots that can be directly (or near-directly) controlled include wearable robotics (e.g., exoskeletons and rehab robotics) and surgical robotics (e.g., the da Vinci surgical system). Wearable robotics have been used in a wide variety of settings, ranging from medical applications (e.g., assistive robotics for rehab purposes) to exoskeletal devices to support physical demand and task performance (e.g., healthcare, agricultural, military, and industrial settings such as construction) (Cha et al., 2020; de Looze et al., 2016; Kim et al., 2019). Surgical robotic systems such as the da Vinci typically consist of surgical robotic arms and a console for the surgeon to control the system via foot pedals, master controls, and controls to adjust positioning (Wu et al., 2020). Catchpole et al. (2019) identified a number of challenges associated with robotic surgeries: postural stress relocated rather than reduced for the surgeon; workflow disruptions which have training implications; communications (verbal and non-verbal) within the surgical team fundamentally affected due to the surgeon being away from the operating table.

Within Line of Sight or Short-Range Teleoperation Within line-of-sight and short-range teleoperations have been applied to the following areas: photography and cinematography (e.g., filming for drama, music, or sporting events), disaster responses and search and rescue activities (e.g., September 11th, Fukushima nuclear power plant, mine and oil disasters, firefighting, earthquakes, hurricanes, floods, and other natural disasters, see Murphy (2014) *Disaster Robotics*), environment inspections and treatments (e.g., warehouse inspection, agricultural applications such as crop inspection, application of chemicals, soil and field analysis), military missions or law enforcement (e.g., aerial surveillance & reconnaissance, hazardous materials and explosives detection and disposal) (Barnes, Elliott, et al., 2019; Oron-Gilad & Parmet, 2017; Pettitt, Redden, Carstens, & Hooper, 2012), and space applications (e.g., astronauts working with robotic systems on the International Space Station (Fong et al., 2013)). Thorough reviews of human performance issues and user interface designs are available in Chen, Haas, and Barnes (2007) and *Disaster Robotics* (Murphy, 2014).

Long-Range Teleoperations Long-range teleoperations are conducted in a wide variety of environments: from telemedicine, to undersea, to space (Fong et al., 2013; Khan & Anwar, 2020; Sivčev et al., 2018). Telemedicine applications include examinations and surgeries via telerobotic systems such as the da Vinci system (Giuliani et al., 2020; Khan & Anwar, 2020; Laaki, Miche, & Tammi, 2019; Valner, Kruusamäe, & Pryor, 2018). Underwater telerobotic systems have been used in a wide variety of environments for diverse purposes such as offshore oil and gas, salvage of sunken objects, biological and geological sampling, and archaeological research, to name just a few (Sivčev et al., 2018). In space-related teleoperations, Earth-based ground control personnel typically operate the remote (e.g., on Mars) robot via preplanned command sequences due to the constraints associated with communications latency (as long as tens of minutes). In the case of the International Space Station, the Earth-based ground control team often work alongside astronauts to operate robotic systems (e.g., the Mobile Servicing System on board the International Space Station). There are several crew performance issues associated with these types of operations: time scale differences (long latency for ground control team but real time for astronauts), control authority coordination and sharing, action planning, and information transfer (Fong et al., 2013). These teleoperations often have to deal with data communication issues (e.g., latency, bandwidth, and availability) that are also challenges for space teleoperations.

2.1.2 Common HRI Techniques

Manual Controllers Manually operated robots can be operated through a wide variety of control media (see Young & Peschel, 2020, for a detailed review), ranging from hand-held devices such as cellular phones (Pitman & Cummings, 2012), game controllers/joysticks (Bonaiuto et al., 2017), gesture-based control devices (Ibrahimović et al., 2019), to augmented reality-based capabilities (Walker, Hedayati, & Szafir, 2019). For manual control and teleoperations, joysticks and game controllers are widely used to operate robotic systems in a wide range of environments (e.g., ground, air, and underwater). See Figure 1 for an example of an indoor mini HRI testbed (Cooke, Demir, & Huang, 2020). To ensure the safety of the operations (e.g., collision avoidance), shared control (human-directed guarded motion) and safeguard mechanisms are often adopted (e.g., nudging, Pitman & Cummings, 2012). In the case of non-line-of-sight teleoperations, similar assisted strategies and techniques are often adopted to ensure safe navigation and operations. For example, the robot's sensors can be used to estimate its poses (position and orientation) for collision and obstacle avoidance purposes, given input from the human operator's commands. The human operator's situation awareness of the robot's environment can be augmented by streaming information from the robot including the red-green-blue (RGB) and depth images, status of the robot (e.g., battery level, CPU load, and other system health information), and 3D point cloud representation of potential obstacles in the robot's immediate path (Perez-Grau et al., 2017).

In situations where the robotic system is located remotely (e.g., space or undersea) and there are significant time delays involving robot control and manipulations, teleoperations often employ mitigation strategies to deal with latency-related issues. For example, in space teleoperations from Earth, in which case latency may be as long as tens of minutes, command sequencing is more feasible than direct manual control (Fong et al., 2013). In environments where telemanipulation needs to be conducted through robotic arms and manipulators/grippers, master/slave controllers with force feedback are often employed to minimize the risk of accidental damage to the manipulator or objects of interest (Sivčev et al., 2018).

Gesture and Haptic Control An intuitive way to control a robot is through gesture control, which enables the human

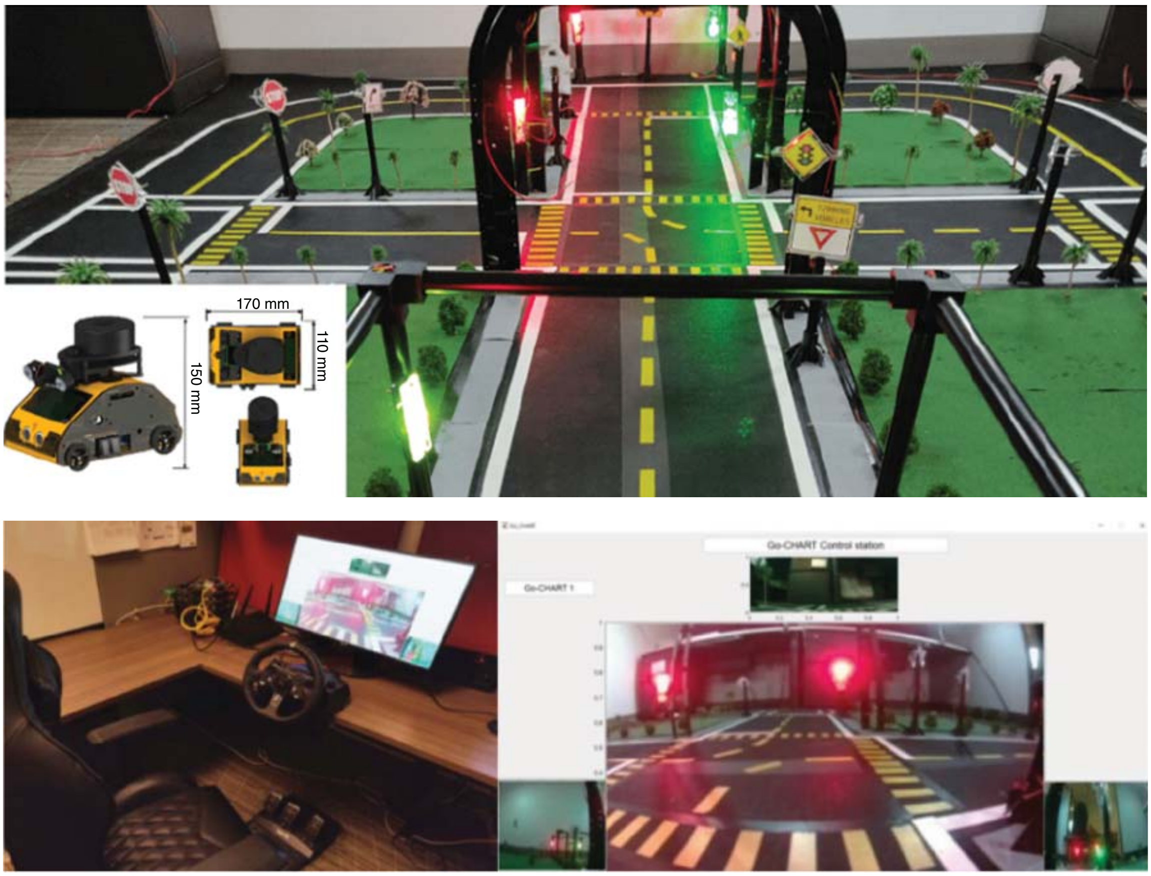

Figure 1 CHARTopolis mini testbed at the Arizona State University (top); human participant experiment station and participant screen view of the testbed (bottom). (*Source:* Courtesy of Shenbagaraj Kannapiran and Nancy Cooke of the Arizona State University.)

operator to convey commands through gestures (e.g., hand or arm). For example, Che et al. (2018) designed a hand-wearable device that can sense the tilting angles of the hand and send commands to robot. Another common gesture control technique is accomplished through haptic gloves (Hansberger 2019; Ibrahimović et al., 2019). A detailed discussion on gesture and haptic control will be presented in Section 3.4.4.

Virtual, Augmented, & Mixed Reality (VAMR) In recent years, there have been increasing applications of Virtual, Augmented, & Mixed Reality (VAMR) capabilities to HRI (this section will only summarize VAMR techniques that are used for teleoperation; the applications for interaction with autonomous robots will be presented in Section 2.2) (Rosen et al., 2019; Szafir, 2019; Williams, Szafir, & Chakraborti, 2019). For local teleoperation, augmented reality (AR) head-mounted displays (HMD) are utilized to augment information (e.g., related to the robot, the environment, or the HRI) available to the human operator, who can view the robot in its mission environment (see Figure 2 for an example of a drone teleoperation experimental set-up; Walker et al., 2019). For example, interface designs can convey to the human operator what is within the robot's field of view (e.g., live camera feed from the robot, virtual imagery overlaid onto the environment) and support communications with the robot by displaying virtual objects/imagery to the operator via the HMD (Walker et al., 2019). For remote teleoperation, virtual reality (VR) HMD can be used to enable human interaction with a remote robot via an avatar, which serves as a virtual surrogate and enables the operator to preview results of different inputs and make corrections if necessary (Ibrahimović et al., 2019; Walker et al., 2019). VR HMD-based remote teleoperation can also be accomplished by mapping the human operator's head movements to the attitude of the robot to support egocentric views from the robot (Zhao, Allison, Vinnikov, & Jennings, 2018). However, latency issues (due to vehicle dynamics, network transmission, computation, etc.) associated with delayed visual updates after the vehicle receives the control commands pose human performance challenges and may cause sickness in some individuals. VR technologies have also been applied to space settings to enable astronauts on board the International Space Station to control a humanoid robot Robonaut 2 (R2) via a VR HMD to perform tasks that require dexterity (Fong et al., 2013).

2.2 Single Robot: Semiautonomous and Autonomous

With the advances in autonomy and artificial intelligence, robotic systems are increasingly able to perform their tasks semiautonomously or autonomously. These robots, humanoid or otherwise, are used in a wide variety of settings—from domestic to undersea to space. Frequently, human operators intervene or teleoperate the robot when human input is required. For example, in the areas of agricultural robotics, human input may be required for tasks such as target detection/selection for harvesting (Vasconez, Kantor, & Auat Cheein, 2019). With higher levels of autonomy, the human operators no longer have to be constantly "*in* the loop" (i.e., active control), rather, they only need to be "*on* the loop." However, staying "on the loop" does not make operators' jobs easier; as Parasuraman and Manzey (2010) noted, automation does not simply perform tasks for humans—it changes the nature of humans' tasks. There are several potential human performance issues associated with increased autonomy: tunnel vision, degraded situation awareness (SA), misuse and disuse of automated systems, and complacency (Chen & Barnes, 2014; Chen et al., 2011; Parasuraman & Riley, 1997). These human performance issues will be examined in detail in Section 4.

2.2.1 Applications

Non-Humanoid Robots There are several categories of non-humanoid autonomous robots: industrial robotics (e.g., in manufacturing, warehouse logistics, agriculture, construction, and mining), domestic/assistive robots (e.g., robots that help with household chores such as Roomba, robotic helpers in hospitality and retail settings), self-navigating robotic vehicles (e.g., ground, air, and underwater), and space robotics (e.g., NASA's planetary rovers and Astrobee). For mobile robots that navigate in their tasking environments, HRI may involve planning (e.g., route planning), which may be done autonomously by the robot or pre-planned by the human operator (e.g., waypoint), or accomplished in a collaborative fashion by both parties (e.g., mixed-initiative, Jiang & Arkin, 2015). For plans that are generated by the robot, transparency (e.g., robot's reasoning process, projected outcomes, etc.) is an important issue that needs to be incorporated into the user interface (see Section 3 on Human-Robot Communications; Chen et al., 2018).

Humanoid Robots Humanoid robots, with various levels of human-likeness—ranging from android (Ishiguro's geminoid) to desktop devices (e.g., Jibo)—typically sense their environments with their eyes, manipulate objects with their arms and hands, and/or move in their environments with human-like motions (Sheridan, 1992, p. 5). Although there are humanoid robots that can be teleoperated, most of these robots perform with certain level of autonomy. Many applications of remotely operated humanoid robots involve therapeutic and rehabilitative practices (e.g., physical therapy, social therapy for autistic children, and elderly assistance) (Goodrich, Crandall, & Barakova, 2013). In some cases, the remote therapist/teacher may interact with the patient/student via a learning from demonstration (see HRI Techniques below) or Wizard of Oz paradigm (Riek, 2012). Fritsche et al. (2015) describe a low-cost multimodal

(a) (b) (c) (d)

Figure 2 (a) Human participant in a VAMR-based drone teleoperation experiment; (b) Experiment environment; (c) Point of interest/target cylinder and scanning progress bar; (d) Teleoperation handheld controller and control mapping: (1) rotation and vertical translation; (2) confirm waypoint placement; (3) horizontal translation; (4) toggle pause/resume. (*Source:* Courtesy of Daniel Szafir of University of Colorado Boulder).

HRI paradigm of humanoid teleoperation using a haptic glove and commercially available HMD (Oculus Rift) and motion tracker (Microsoft Kinect). One high-profile humanoid robot is NASA's Robonaut 2 (R2), which is the first humanoid robot designed specifically for space and has been deployed on the International Space Station since 2011. R2 serves as an astronaut assistant and works side by side with astronauts in repair and maintenance tasks and other housekeeping chores (Fong et al., 2013). R2 can be teleoperated by astronauts on board the ISS via a telepresence system (VR HMD); R2 can also be remotely operated by ground-based controllers in Houston, Texas, via scripted command sequences.

2.2.2 Common HRI Techniques

Waypoint Operation Waypoint operation is a common way for the human operator to provide navigational guidance to the robot while enabling the robot's onboard autonomy (e.g., collision avoidance) to traverse in the environment. Along with waypoints, other robot- and task-related information (e.g., robot's heading, field of view, and velocity, areas of interest and landmarks) are also often displayed on the map. When more fine-grained control of the robot is required (e.g., when approaching an area of interest), techniques such as nudge control (Pitman & Cummings, 2012) are useful to allow the operator to provide more detailed guidance to the robot (e.g., vehicle altitude, camera viewpoint), which then completes the task autonomously.

Learning from Human Demonstration With the learning from demonstration paradigm, the human operator can teach a robot how to complete a task by showing the robot the sequence of the task (Chernova & Thomaz, 2014; Fitzgerald, Goel, & Thomaz, 2018). There are multiple ways this can be accomplished. The operator can physically move the robot's joints to indicate the motion sequence, which the robot records and then segments into task steps (e.g., opening or closing its gripper). These motion sequences can also be captured directly from the human through motion capture and gesture-based programming. The human operator can also provide instructions (e.g., trajectories or keyframes of the robot's joints) through a computer-based interface by manipulating a 3D representation of the robot.

Natural Language Interaction Due to the advances in AI in recent years, applications of natural language processing (including both understanding and generation) to HRI have become increasingly sophisticated (Briggs, Williams, & Scheutz, 2017; Frasca, Oosterveld, Krause, & Scheutz, 2018; St Clair & Mataric, 2015; Veloso, 2018). These spoken dialog systems are widely used in consumer products in mobile phones (e.g., Siri) and smart speaker platforms (e.g., Alexa, Google Home). In the context of HRI, natural language capabilities have been applied to enable a natural and intuitive way for humans to interact with robots in not only well-controlled laboratories but also the real world (see "Situated interaction," Bohus & Horvitz, 2019; "Robots in the wild," Jung & Hinds, 2018). More on natural language interaction will be presented in Section 3, Human-Robot Communications.

2.3 Multiple Robots and Swarms

Due to progress in AI and autonomy-related fields, it has been increasingly feasible for human operators to manage more than one robot at a time, from marsupial robot systems to multirobot teams to swarms. These robotic systems are deployed in such diverse settings as construction, search and rescue and disaster management, military missions (see Figure 3 for an example of a simulation testbed at U.S. Air Force Research Laboratory; Calhoun et al., 2018), firefighting, nuclear operations, bomb squad and landmine detection, warehouses, agriculture, coastal patrol, and underwater operations (Draper et al., 2018; Lewis, 2013; Murphy, 2014; Rosenfeld, Agmon, Maksimov, & Kraus, 2017; Vasconez et al., 2019). There are several potential human performance issues associated with multirobot control: span of control and multitasking; task switching (among vehicles and among different types of tasks, e.g., navigation and target detection); situation awareness, and change blindness (Chen & Barnes, 2014; Chen et al., 2011). These human performance issues will be examined in detail in Section 4.

2.3.1 Applications

Visual Assistant and Marsupial Systems As discussed in the teleoperation section, a major human performance issue is degraded remote perception, particularly limited field of view. One of the mitigation strategies to deal with this challenge is to use one more robot (homogeneous or otherwise) to serve as a visual assistant (Xiao et al., in press; see Figure 4). For example, pairs of homogeneous robots were used in the Fukushima Daiichi nuclear power plant response and the Deepwater Horizon Spill in order to enhance task performance and to save time (Murphy, 2014). More frequently, though, heterogeneous robots (typically one being an aerial vehicle) are used so one platform can compensate for the other's limitations. For example, a drone (tethered to the other platform or not) is often paired with a ground or surface robotic vehicle to provide an aerial view (Kalaitzakis et al., 2020; Kiribayashi, Yakushigawa, & Nagatani, 2018; Xiao et al. in press). In some cases, these pairs/groups of robots are operated as a so-called "marsupial"

Figure 3 IMPACT testbed for multirobot management research (Center top: real-time tactical map; Center bottom: sandbox and Playbook interfaces; Right: sensor feeds from unmanned vehicles). (*Source:* Courtesy of Gloria Calhoun of U.S. Air Force Research Laboratory.)

Figure 4 Visual assistant interface (right) and PackBot uPoint controller interface (left). (*Source:* Courtesy of Robin Murphy of Texas A&M University.)

system, meaning that the larger vehicle can carry the smaller vehicle(s) to the operational site (Stankiewicz et al., 2018).

Multirobot Control Multirobot systems can be classified based on their size (i.e., number of robots in the system), communication aspects (e.g., bandwidth, range, and topology), reconfigurability, processing ability, and composition/heterogeneity (Dudek, Jenkin, & Milios, 2002). Due to the nature of multirobot systems (e.g., availability of multiple robots simultaneously and their potential abilities to be spatially distributed in the tasking environment), these systems are often used for foraging purposes, such as search and rescue activities (Lewis, 2013) and other exploration and reconnaissance tasks (Draper et al., 2018; Mercado et al., 2016; Stowers et al., 2020; Vered et al., 2020). Related to human span of control, fan-out (the number of robots a human operator can control simultaneously) models have been proposed that take into account the *activity time* (the time that a robot is active) and *interaction time* (the time that it takes for a human operator to interact with a robot), and *neglect tolerance* (how much time a robot can be neglected by a human operator without performance degradation) (Crandall et al., 2005; Olsen & Wood, 2004). Based on a review performed by Lewis (2013), human operators typically attempt to control as many robots as available, and the decline of human operator's performance per robot tends to be gradual (rather than an abrupt drop-off at a certain limit), as the number of robots increases. Although Wang et al. (2009) suggest that the maximum number of robots a human operator can operate simultaneously is between four and nine plus robots, depending on the robot's autonomy level and environmental factors, multiple studies observe degradation of human operator performance when the number of robots under human control increases above four (Chen & Barnes, 2012b; Chien et al., 2013; Rosenfeld et al., 2017). Typically, multirobot systems are managed by humans via supervisory control (see Chapter 28 by Sheridan in this volume), often involving co-located or distributed teams (see an example of a simulation testbed for a distributed multirobot management team). Some of the multirobot control techniques (e.g., supervisory control via automated planning, Playbook) are presented in the following Section 2.3.2; more information about multimodal (speech, gesture, and haptic) control is available in Section 3.4.

Swarms Robotic swarms refer to large number (typically > 50) of (usually homogeneous and miniature) robotic agents that coordinate with one another via simple local communications and control laws (e.g., attraction, repulsion, and orientation) to exhibit emergent behaviors (e.g., flocking, rendezvous, and dispersion) (Brown et al., 2016; Chandarana, 2019; Couzin et al., 2002; Crandall et al., 2017; Nam et al., 2020; Roundtree et al., 2020). While many robotic swarms are based on bioinspired paradigms (e.g., flocks of birds, schools of fish, and bee colonies), there are other approaches such as organizational/social, knowledge-based, and ontological/semantic paradigms (Kolling et al., 2016). Swarm technologies have been applied to a wide range of operational settings, such as search and rescue, military surveillance and reconnaissance, environmental monitoring and tracking, agriculture, and space exploration (Kolling et al., 2016). Although each agent has simple and limited capabilities, collectively, swarms can achieve substantial robustness and resiliency due to the responsibilities being distributed among the members of the swarm. Swarms also support flexibility/scalability and can adopt different group sizes and behaviors based on environmental influences or human input (see Figure 5 for an example of a human-swarm interface for shared control; Crandall et al., 2017). However, these attributes along with constraints typically associated with swarms (e.g., limited sensing and communications capabilities, latency and asynchrony between human input and swarm responses) can pose challenges to human-swarm interaction effectiveness (e.g., assessing the state and dynamics of the swarm, predicting emergent behaviors based on human input) (see Roundtree, Goodrich, & Adams, 2019, for challenges associated with applying transparency design principles based on non-swarm interfaces). For reviews of human-swarm interface designs (e.g., transparent interfaces, predictive displays and visualizations, and speech interfaces), see Hocraffer and Nam (2017), Kolling et al. (2016) and Roundtree et al. (2020).

2.3.2 HRI Techniques

Supervisory Control of Multirobot Systems For supervisory control of multiple individual robots, two paradigms have been studied extensively: management by consent (MBC) and management by exception (MBE). MBC requires the robot to ask for explicit consent from the human operator before taking any actions; MBE, on the other hand, allows the robot to initiate actions unless overruled by the human operator. Overall, research findings (see Chen et al., 2011, for a detailed review) suggest that the benefits of MBE versus MBC paradigms for supervisory control of multirobot systems are not always consistent and they are affected by factors such as the difficulty level of the task and the type of automation unreliability. Additionally, operators may exhibit complacency under high task load when using systems with higher levels of autonomy. Mitigation strategies such as adjustable or adaptive automation have been proposed to address weaknesses of fixed automation schemes such as MBC and MBE. A thorough review is available in Chen and Barnes (2014).

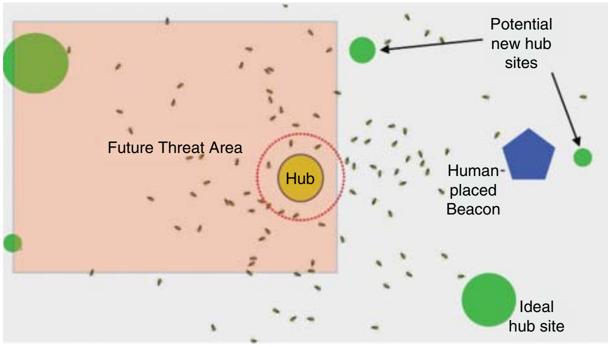

Figure 5 Shared control for human-swarm interaction (the swarm finds potential hub sites based on human guidance; larger green circles indicate better options). (*Source:* Courtesy of Jacob Crandall and Michael Goodrich of Brigham Young University.)

Supervisory Control via Automated Planners Given the advances in intelligent planning capabilities and the limitations of human span of control, research has been conducted to investigate the effectiveness of using an intelligent agent to assist the human operator with managing multiple robots (Chen & Barnes, 2012a, 2012b; Draper et al., 2018; Miller & Parasuraman, 2007; Rosenfeld et al., 2017). For example, RoboLeader (Chen & Barnes, 2012a, 2012b) was developed to assist the human operator with collecting information from subordinate robots with limited autonomy (e.g., collision avoidance and self-guidance to reach target locations), making tactical decisions, and coordinating the robots by issuing commands, waypoints, or motion trajectories. In typical mission situations, RoboLeader would recommend route revisions when encountering environmental events that require robots to be rerouted. The human operators, in turn, can accept the plan revisions or modify them as appropriate. A similar paradigm to RoboLeader is reported in Rosenfeld et al. (2017) that can advise the human operator on multirobot management in complex tasking environments (e.g., search and rescue and warehouse operations). The advising agent capabilities have been tested using both simulated and physical robots, and the results show that the human operators' performance was significantly improved when working with the advising agent to manage a team of ten robots. Similarly, RoboLeader was effective in enhancing the human-autonomy team performance while reducing the operators' workload; however, effects of operator individual differences consistently impacted the operator performance (Chen & Barnes, 2012a, 2012b; Wright, Chen., & Barnes, 2018). Another example of an intelligent planning agent for managing multiple robots is the *Playbook* system, which uses the American football analogy of play-calling to task the agents (Miller & Parasuraman, 2007). A major advantage of such a planning agent is the flexibility available to the human operator, who can choose *plays* from the existing Playbook, work with agent-recommended plays and tweak them if necessary, and delegate detailed planning tasks to the agent. The Playbook system has been investigated extensively, and the latest studies are reported in Draper et al. (2018) and Calhoun et al. (2018).

3 HUMAN–ROBOT COMMUNICATIONS

As robots evolve from tools to human teammates, the dynamics of their interactions also change dramatically. The type of human-robot communications can take many different forms, from traditional human-computer interaction to language-based and multimodal interaction. This section reviews common human-robot communications and their associated research.

3.1 Language-Based Interaction

Human ability to use language allows them to understand and communicate abstract concepts and even emotional context. A common linguistic framework is essential for joint problem-solving and for communicating intent in complex situations. The structure of language is itself complex, consisting of individual symbols (morphology), semantics (their meaning), syntax (rules for sentences), and pragmatics (context) (Jurafsky & Martin, 2009). Usually (although not always), human teams not only share a common language but also understand the context of verbal or other symbolic representations necessary to communicate their intentions (for example, a military mission). Human-agent communications are more complicated for a number of reasons: ambiguous semantics, unclear intent, and variable context.

Natural language processing (NLP) is a branch of artificial intelligence (AI) whose aim is to interpret and generate natural language from text, verbal utterances or other media enabling discourse between humans and machines (Jurafsky & Martin, 2009). NLP is still a developing discipline but recent advances promise to make the technology more amenable to creating useful interactions between humans and machines in areas such as two-way communications, the ability of agents to initiate the dialogue, agent interpretation of human emotional states, multimodal communications, and linguistic interpretation of spatial representations (Evans et al., 2017; Mavridis, 2015; Skubic et al., 2004; Tellex et al., 2011). The major obstacle for wider use of open-ended NLP systems is the complexity inherent in the real world as well as linguistic ambiguities that make translation difficult (Jurafsky & Martin, 2009). Fortunately, the importance of NLP for commercial use is accelerating the pace of progress.

To be practical in high-stake tasking environments, such as military operations and disaster relief, NLP must be able to interpret multimodal inputs (see Section 3.4) including text and verbal utterances which then must be processed by an agent's cognitive layer as inputs or requests related to solving a particular problem (Barnes, Lakhmani, et al., 2019). The cognitive layer could consist of any number of algorithmic or AI approaches, depending on the type of problems the agent is expected to solve (Barnes, Chen, & Hill, 2017). Machine learning (ML) approaches based on reinforcement learning are currently being used for complex problem sets because of the ML ability to solve problems that were previously considered intractable (Everitt & Hunter, 2018). However, although ML solutions have proven efficient for diverse problem sets, its underlying processing is often opaque (Chakraborty et al., 2017). To address the opaqueness quandary, researchers are investigating eXplainable AI (XAI) techniques to decipher

ML and use transparency techniques (see Section 3.2) to help operators visualize ML intent and logical underpinnings (Mueller et al., 2019; Pynadath, Barnes, et al., 2018).

3.2 Agent Transparency

Agent software architectures should be designed to complement their human partners not only for the tasks that agents are assigned but so that agents also have some degree of mutual transparency with their human partner (Lyons, 2013). Chen and her colleagues defined agent transparency as "descriptive quality of an interface pertaining to its abilities to afford an operator's comprehension about an intelligent agent's intent, performance, future plans, and reasoning process" (Chen, Procci, et al., 2014, p. 2). Empirical evidence from a variety of studies (Chen et al., 2018; Lee, 2012; Lyons & Havig, 2014; Schaefer et al., 2017) indicates that increasing agent transparency improves both human's trust in the agent and the overall human-agent team performance (see a recent special issue on Agent and System Transparency published in *IEEE Transactions on Human-Machine Systems* for the latest research findings (Chen et al., 2020)). The efficacy of a specific transparency model depends on the task type, the task load and how transparency is implemented (Bhaskara et al., 2020; Wright, Chen, & Lakhmani, 2020). A distinction should be made between transparency-related trust (discussed in detail in Section 4) and calibrated trust (reduction of disuse and misuse). Transparency-related trust depends on a number of factors including individual differences, cultural background, previous experience with autonomous systems, physical configuration, and the robot's perceived reliability (Bhaskara et al., 2020; Chien et al., 2020; Dzindolet et al., 2003; Matthews et al., 2020; Nam et al., 2020; Schaefer, Hill & Jentsch, 2018).

One of the transparency frameworks that has been developed and investigated in a variety of simulation environments is Situation awareness-based Agent Transparency (SAT) (Bhaskara et al., 2020; Chen et al., 2018; Roth et al., 2020; Stowers et al., 2020; Wright et al., 2020), which is based on both Lee's (2012) model incorporating the 3Ps of the system (purpose, process, and performance) and Endsley's (1995) model of situation awareness (SA). SAT posits that a transparent agent should support the human operator's SA at all three levels: Level 1 (L1), perception of its actions and plans; Level 2 (L2), comprehension of its logic; and Level 3 (L3), prediction of intended end-state including uncertainties (U) (Chen et al., 2018). SAT was evaluated as part of the U.S. Department of Defense (DoD) Autonomous Research Pilot Initiative program for both a planning agent in the context of multirobot management (the IMPACT project; Mercado et al., 2016; Stowers et al., 2020; Figure 6) and a small ground robot application (the Autonomous Squad Member, ASM project; Selkowitz, Lakhmani, & Chen, 2017; Wright et al., 2020; Figure 7). In general, correct usage (i.e., calibrated trust) and SA increased for higher levels of transparency based on SAT—L1<L12<L123 and L123+U in both paradigms. Although the trends were consistent, each increment was not always statistically significant across studies. Subjective trust generally followed this trend but the presence of uncertainty-related information or the system's perceived reliability could affect trust assessment (Mercado et al., 2016; Selkowitz et al., 2017; Stowers et al., 2020; Wright et al., 2020). Giving operators uncertainty information (e.g., poor weather may cause delays) improved appropriate reliance/compliance but also increased operator response latency by a few seconds (Stowers et al., 2020). Kunze et al. (2019) found that, during human interaction with automated driving, transparency information conveying uncertainties benefited the operator's overall

Figure 6 IMPACT simulation experimental environment. Transparency designs based on SAT: "projected plan success" (Mission Metrics in lower left) (unfilled circles convey uncertainty), "plan specifications" text box (orange text conveys uncertainty associated with reasoning), and the map areas (vehicles/items with uncertain status/info are translucent). (*Source:* Courtesy of U.S. Army Research Laboratory.)

Figure 7 Autonomous Squad Member human-robot interface. "At-a-glance" transparency module based on the SAT framework is shown in the upper left corner. (*Source:* Courtesy of U.S. Army Research Laboratory.)

Figure 9 Transparent human-swarm interfaces with collective visualization (top) or individual-agent visualization (bottom). (*Source:* Courtesy of Julie Adams and Karina Roundtree of Oregon State University.)

task performance (including trust calibration and situation awareness) but significantly increased the operator's workload. Taken together, these studies indicate that uncertainty could have a positive effect on task performance but a negative effect on operator workload and response time; a tradeoff that makes sense for complex planning missions such the IMPACT missions (Calhoun et al., 2018; Draper et al., 2018) but uncertainty information may not be useful for ASM-type missions that require instantaneous decision making (Selkowitz et al., 2017; Wright et al., 2020).

A SAT-based interface was implemented for a workload-adaptive cognitive agent in a multirobot management context to support a helicopter crew (Roth et al., 2020; Figure 8) and a human–swarm interface (Roundtree et al., 2019; Roundtree et al., 2020; Figure 9). Roth et al. (2020) found that, similar to the findings reported above, transparency significantly enhanced the human teammates' SA and task performance. Additionally, the transparent agent was perceived as more humanlike than the opaque agent. Roundtree et al. (2019) provide a detailed analysis of the challenges of designing transparent human-swarm interfaces to achieve the three levels of SAT, and in a follow-on paper, Roundtree et al. (2020) report an experiment comparing two different visualization design methods (collective vs. individual agent; see Figure 9). Results showed that the collective design achieved better transparency and supported human participants' SA better than the individual agent design. It is worth noting that the utility of SAT (or any transparency model) depends on the effectiveness of how the information is presented and not just on what information is presented (see Section 3.4). Efficacy depends on following human factors principles, effective designs (e.g., visualizations) for a specific environment, and experimental verifications (J. Cha et al., 2019). Recent efforts have also started to examine ways to promote bidirectional transparency—not only of the robot (to the human) but also of the human (to the robot) (Chen et al., 2018; Lyons, 2013; Wynne & Lyons, 2018). To foster mutual transparency, interaction between humans and agents needs to be an ongoing process. However, developing effective communication protocols between humans and agents in an ever-changing environment is a difficult challenge. Finally, the findings reported in Matthews et al. (2020) suggest that transparency content should be compatible with the operator's mental model and, when possible, personalized to highlight the robot's data-analytic capabilities or its social (e.g., humanlike) aspects.

Figure 8 Transparent workload-adaptive multirobot management interface. (*Source:* Courtesy of Axel Schulte and Gunar Roth of University of the Bundeswehr Munich.)

3.3 Implicit Communications

As Cooke (2015) has stressed, a shared knowledge base is not sufficient for team interactions; each member must understand when to push knowledge or request it from their partners to ensure efficient collaboration. Efficient collaboration also assumes that team members can anticipate and react to each other's actions without overt communications (Salas et al., 2015). Human teams, especially experienced ones, are able to act in concert because humans have evolved the ability referred to as a *Theory of Mind* (ToM) that enables them to have insight into each other's mental states, expected future requirements, and even emotional content—a process which is strengthened through continual interactions (Mahey, Moses, & Pfiefer, 2014). For HRI, researchers have used reinforcement learning to train robots to anticipate and to act (Pynadath & Marsella, 2005). Pynadath and Marsella (2005) developed a recursive simulation environment (*PsychSim*) to emulate ToM of various agents by developing a software representation of a shared mental model among the agents. Understanding of characteristics attributed to each agent were refined by recursive simulations enabling agents to respond appropriately to other agents based on their perception of that agent's behavioral profile.

In a series of studies using *PsychSim*, Wang and her colleagues investigated agent transparency (based on the SAT framework) in the context of human-robot teaming for military tasks (e.g., reconnaissance). Results show that humans trusted and complied with higher performing robots; however, if poorer performing robots explained the reason for its decision, its human teammate was able to use the robot's explanation to know when to override incorrect suggestions minimizing performance differences between robots of differing reliability level (Wang et al., 2016). The results imply that overt communication is essential for human-agent teaming with less proficient robots whereas compliance with more effective robots depends on human perception of their past performance (Hancock et al., 2011). Further research indicates that operator knowledge of the robot's confidence level improved participants' compliance, trust, and correct safety decisions, again, especially in identifying errors. Current *PsychSim* applications make human-agent cooperation more efficient but humans still need bidirectional communications to verify the robot's decisions (Barnes et al., in press; Pynadath, Wang, et al., 2018; Wang et al., 2018). Richer bidirectional and multimodal communications (e.g., incorporating visualizations based on the SAT) may not only create a vehicle for more efficient communications but also aid in fostering shared mental models for humans and agents to act cooperatively based on mutual understanding of the tasking environment (McNeese et al., 2017; Pynadath, Barnes, et al., 2018).

3.4 Multimodal Interfaces

As robots become more prevalent, their complexity and possible uses also increase exponentially. Robotic systems frequently have multiple inputs and subsystems that require constant monitoring not only by the onboard systems but also by human operators to supervise and coordinate systems per the overall developing situation (Defense Science Board, 2016). Maintaining the SA of a volatile environment such as urban traffic while monitoring one or more autonomous systems presents significant challenges and tests the limits of human processing capacity (Gillan, Riley, & McDermott, 2010). Just as humans use multiple sensory inputs (visual, auditory, tactile, vestibular, etc.) to maintain SA during quotidian interactions; information from multiple sensory modalities has been shown to improve operator processing efficiency (Wickens, 2008). Multimodal interfaces not only unburden the operator, they also enable humans to interact with their environment in a more naturalistic manner (Elliott & Redden, 2013; Wickens et al., 2011).

Multiple sensory channels are more than additional sources of information; they combine to give humans a more holistic understanding of an ever-changing milieu of inputs. Designing multimodal interfaces has two criteria: (1) designing effective multimodal solutions for a particular task; and (2) designing synergistic multimodal combinations for multitasking environments (Elliott & Redden, 2013). Young and Peschel (2020) conducted an extensive review of interfaces for small robotic systems (aerial, ground, and sub-surface), discussing the advantages and limitations of various approaches, stressing the importance of interfaces that are designed so controllers (e.g., haptic) are combined with displays (e.g., stereoscopic) optimized for specific tasks (e.g., disarming mines) within the context of interface limitations (e.g., bandwidth). Certain interfaces (master-controller) are designed for specific tasks such as using a replica of a robotic arm to control a remote robotic arm. Young and Peschel (2020) suggest that as robots become autonomous, VAMR (see Section 2.1.2) is a promising solution to enable operators to supervise systems in complex environments. These enhancements will improve the quality of telepresence, especially in difficult environments such as sub-surface nautical applications. The following is a brief introduction to the advantages and limitations of multimodal interfaces, many of which were tested in diverse real-world environments.

3.4.1 Visual Displays

The principal advantage of conventional visual displays is their ability to mix text, graphics, and imagery from multiple sources of data (J. Cha et al., 2019; Ophir-Arbelle et al., 2013; Oron-Gilad, 2014). Augmented reality displays enhance real-world tasks such as guiding unmanned aerial vehicle (UAV) operators through dangerous terrain by using both synthetic and sensor imagery (Calhoun & Draper, 2010). Moreover, recent approaches make the robot itself a rich source of visual information (E. Cha et al., 2016). Eye gaze, an important component of human interaction, has been shown to enhance human-robot bidirectional communication (Admoni & Scassellati, 2017). Cues from changes in an autonomous UAV flight patterns or light signals from aerial platforms can be used to inform ground operators of state changes (Szafir, Mutlu, & Fong, 2017). Other visual interfaces such as stereoscopic displays has been investigated in simulation-based HRI experiments (Chen, Oden, & Merritt, 2014) and tested in field exercises for disarming simulated explosive devices with a robotic arm (Bodenhammer, 2007). However, the richness of visual information is also their major flaw, visual displays can easily overwhelm operators with a profusion of data especially for multitasking environments causing the operator to focus on a single task. During a high workload, cognitive efficiency has been shown to improve by combining vision with other sensory modalities. Multimodal information helps to overcome cognitive tunneling in the visual domain by using other sensory channels to improve overall SA (Elliott, Coovert, & Redden, 2009; Wickens, 2008).

3.4.2 Speech and Auditory Interfaces

Burke et al. (2006) conducted a meta-analysis of 24 studies indicating that a combination of auditory and visual displays resulted in better performance than either modality in isolation. Soldiers in a combat environment rely on auditory cues and radio communications to maintain SA and communicate their situation to their superiors (Barber et al., 2014). 3D auditory displays have the capability to redirect attention to different locations or different radio streams, making audition an ideal medium to keep track of the current status and location of multiple robots without diverting attention from other displays, although the perceptual ambiguity between frontal and rear (i.e., the cone of confusion) requires disambiguation (Vause et al., 2000). Speech commands derived from soldier lexicons were shown to be a natural hands-free interface for commanding robots during a navigation task (Harris & Barber, 2014). St. Clair and Mataric (2015) demonstrated the utility of speech for two-way communications between humans and autonomous robots working cooperatively in a synthetic task. However, in a multi-UAV management task in a simulated helicopter environment, touch was superior to speech controls, indicating limitations of speech in environments where spatial mapping stereotypes favor automated responses (Levulis, DeLucia, & Kim, 2018).

3.4.3 Haptic Interfaces

In a meta-analysis of 40 scientific articles covering diverse multimodal conditions, tactile displays combined with visual displays are found to be superior compared with visual displays alone (Elliott et al., 2009). Tactile cues put low attentional demand on operators for navigation and tasks related to spatial orientation enabling humans to monitor small robots without interfering with SA (Elliott et al., 2010; Elliott, Skinner, Pettitt, Vice, & Walker, 2014; Redden et al., 2009). Tactor arrays (vests and belts) using signal amplitude, frequency, and location are capable of generating different message sets although tactile communication is still limited compared with auditory or visual channels (Mortimer et al., 2020). For limited message sets, test participants are able to learn 12 different tactile commands in less than 30 minutes and accurately recall meanings several hours after the first training session (Mortimer et al., 2020). Because of their stealth and low bandwidth characteristics (particularly relevant to military applications), tactile interfaces are currently being adapted for bidirectional communication as well as unidirectional robotic control (Barber, Reinerman-Jones, & Matthews, 2015; Scheggi, Aggravi, & Prattichizzo, 2017). Tactile augmentation is efficacious in noisy environments or during vigilance tasks because tactile stimulation redirects attention to mission-critical information when visual and auditory channels are otherwise engaged (Calhoun et al., 2005). Force and velocity cues as well as tactile information are being incorporated into haptic displays to signal change in the status of unmanned vehicles whereas haptic controllers coupled with stereoscope displays are being designed to use force amplitude to control small robots (Edmonson et al., 2012; Pacchierotti et al., 2017; Son et al., 2013). In the medical domain, haptic feedback is not considered a completely satisfactory means to replace the surgeon's natural senses; multimodal applications using tactile, kinesthetic, and vibrotactile feedback are showing improved performance compared to haptic feedback alone (Abiri et al., 2019).

3.4.4 Gesture-Based and Virtual Interfaces

Gestures using either an instrumented glove (Figure 10) or camera-based controllers have been tested successfully to guide robots and to communicate commands based on lexicons of hand gestures currently used by soldiers to coordinate squads during silent maneuvers (Elliott, Hill, & Barnes, 2016). Recent advances have enabled gesture control using a wearable smart watch configuration and as a means to control a quadcopter using hand position on the control grip (Hartnett et al., 2018; Ibrahimović et al., 2019). Pointing using a wrist-mounted inertial mounted measurement unit (IMU) enables operators to point to specific robots in a group of robots, move the robot, receive feedback as to the relative location of the operator and the robot using an intuitive non-verbal interface (Gromov et al., 2019). Gestures can also be used to enrich NLP interfaces by taking advantage of human use of gestures to enhance understanding of spoken dialogue (Meszaros, Chandarana, Trujillo, & Allen, 2018). An innovative use of gesture, eye gaze, and voice controllers has resulted in the creation of virtual worlds that generate large-scale visualizations of stereoscopic representations of operators immersed in real-world scenes. This give a single operator or group of viewers (e.g., Division staff) the capability to explore a real-world scene showing locations of friendly and adversarial forces embedded in realistic geographic terrain. Immersive displays can focus on individual vehicles or groupings of manned and autonomous assets in order to determine their current status permitting the operator to either investigate specific assets or have a "god's eye" view of the visual scene. Gesture control is a natural way for an operator to move in and around 3D immersive displays and has proven effective when compared to traditional gaming interfaces (Coa, Peng, & Hansberger, 2019; Hansberger, 2019; Hansberger et al., 2019; Sanders, Vincenzi, Holley, & Shen, 2019). Although the virtual applications reviewed are being investigated in laboratory settings, current applications of varying degrees of resolution are ubiquitous in the gaming world.

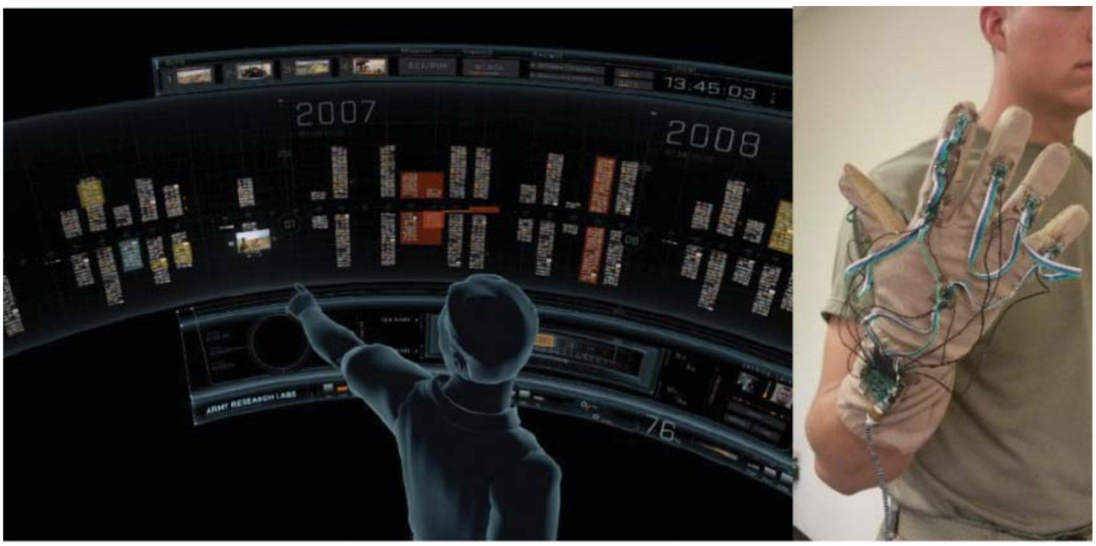

Figure 10 Haptic glove and gesture control for a virtual environment (left) and robotic control (right). (*Source:* Courtesy of Jeffrey Hansberger and Linda Elliott of U.S. Army Combat Capabilities Development Command.)

3.4.5 Anthropomorphism and Proxemics

The embodiment of a robot, its gestures, facial expressions, and speech are the focuses of social robotics (Breazeal, 2002; Karreman, Ludden, & Evers, 2019; Schulz, Torresen, & Herstad, 2019). It is particularly important for service robots in clinical settings to appear benign and professional if they are going to be accepted by elderly or other vulnerable populations (Breazeal et al., 2019). Yogeeswaran et al. (2016) found an interaction between robot appearance and its relative proficiency in tasks usually performed by humans. The dependent measure was their willingness to support robotics research. The participants were least likely to support research for humanoid-appearing robots that were more proficient than humans, suggesting these robots were perceived as threatening in the future. A related phenomenon is the *uncanny valley* effect. Robots that appear almost but not quite human have been shown to produce negative affect in humans. Złotowski et al. (2018) found almost human-appearing robots to have a higher *eeriness* rating and lower likeability rating than a more cartoonish, less human-appearing robot. The less human-appearing robot's likeability rating was especially sensitive to a number of interactions becoming more positive after three sessions of positive interactions and the converse after three sessions of negative interactions.

Social interactions are mediated by a number of factors including eye gaze, size, movements, and characteristics of the robot such as social distancing (Dragan et al., 2015). Proxemics refers to psychological effects of physical space between humans which varies depending on the purpose of the interaction: (1) the intimate distance zone (up to 1.5 feet); (2) the personal distance zone (1.5–4 feet); (3) the social distance zone (4–11.8 feet); and (4) the public distance zone (over 11.8 feet) (Hall, 1966). Humans are threatened not only by the physical distance they are from robots but also the robot's posture (Obaid et al., 2016). Furthermore, rated likability and eye gaze are factors that influence the distance at which humans felt comfortable interacting with a robot (Mumm & Mutlu, 2011). Given that robots and humans will likely work closely together in the future (especially after the CoVID-19 experience), technology is being developed at MIT that is able to sense human-robotic distance in a factory setting to ensure safe proxemics under a variety of working conditions (Lasato, Fong, & Shah, 2014). El-Shawa et al.'s (2017) research emphasized the utility of using virtual robots for training because their participants were able to work more easily at closer distances with virtual robots than with the physical embodiment of the same robot.

3.4.6 Multimodal Synergy

An important advantage of multimodal interfaces is their capability to reinforce or to separate message streams: radio channels to monitor communications and visual channels to view the on-going tactical situation. The ideal mixture of displays depends on understanding the operator's tasking load and mission requirements (Crandall et al., 2005). Tasks requiring coordinated effort should be in close proximity physically and cognitively using modalities that reinforce each other, such as 3D audio alerts and a visual cue on a map to emphasize an immediate emergency (Andre & Wickens, 1992; Haas & van Erp, 2014). Contrasting modalities are useful to split attention between separate tasks in situations where multitasking requires different solutions (e.g., targeting alerts (audio) and teleoperations (visual) (Chen, 2010)). In other tasking environments, dual modalities can be combined to compensate for the disadvantages of either modality in isolation. For example, tactile cueing is better than speech for rapidly assessing direction in a navigation task; speech is better for distance commands (left (tactile), 100 Ft (speech); Elliott, Jensen, Redden, & Pettit, 2012).

4 HUMAN FACTORS ISSUES

4.1 Situation Awareness

Endsley's (1995) now classic definition of situation awareness (SA) states "the perception of elements of the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future." SA is one of the key human performance issues for effective HRI. Human performance issues involved in robot teleoperations generally fall into two categories, i.e., remote perception and remote manipulation (Steinfeld et al., 2006). Teleoperation tends to be challenging because operator performance is "limited by the operator's motor skills and his ability to maintain situation awareness … difficulty building mental models of remote environment … distance estimation and obstacle detection can also be difficult" (Fong, Thorpe, & Baur, 2003, p. 699). In the teleoperating environments, human perception is often compromised because the natural perceptual processing is decoupled from the physical environment. This decoupling affects people's perception of affordances in the remote scene and often creates problems in remote perception such as scale ambiguity and poor depth perception (Sato et al., 2020; Tittle, Roesler, & Woods, 2002). Simple tasks could be challenging due to the lack of motion feedback in remote visual processing as well as the mismatching of viewpoints, which could result from placing the camera at a height that does not match the normal eye height (Tittle et al., 2002). Also related to remote perception challenges is the limited field of view issue. The use of cameras to capture the environment in which the robot is navigating sometimes creates the so-called "keyhole" effect (Voshell, Woods, & Phillips, 2005; Woods et al., 2004). Human operators may be susceptible to "tunneling" effects; depth perception, sense of speed, object detection, and identification of self-location may also be degraded (Darken, Kempster, & Peterson, 2001; Murphy, 2014; Sato et al., 2020; Van Erp & Padmos, 2003). In general, poor perception has a detrimental effect on SA and, therefore, on teleoperating tasks.

Maintaining SA for robotic operations is more complex than understanding the immediate environment in terms of its perceptual and cognitive characteristics; future systems will not operate in isolation or in a static environment. Operators must be aware of the robot as part of the overall tasking environment (e.g., the effects of changing military or traffic patterns), the local situation (e.g., the robot's movements), as well as their combined impact on attaining future objectives. HRI frequently involves multiple concurrent tasks (e.g., planning, monitoring, manipulating, communicating, problem solving, etc.), and the effects of task switching on HRI task performance need to be taken into account during the HRI design process (Gartenberg et al., 2014). Change blindness can also happen when the operator has to manage multiple robots simultaneously (Parasuraman et al., 2009). Empirical evidence showed that operators' loss of SA due to switching attention among robots may require as long as 12 seconds to recover (Goodrich, Quigley, & Cosenzo, 2005). The situation is exacerbated when robots' planning is automatically generated by path planning agents (Marquez & Cummings, 2008) and when robots are highly autonomous and the operator's task is mainly monitoring, which tends to result in vigilance decrements (Wohleber et al., 2019).

The type of information required for autonomy reflects the operator's role, shifting from granular control to focusing on end-states emphasizing both overall SA and SA of the robot's intent in terms of achieving its objectives (Chen et al., 2018; Endsley, 1995; Schaeffer et al., 2017). There is evidence that user interface design frameworks such as the Situation awareness-based Agent Transparency (SAT) can enhance the operator's SA in HRI tasking situations (Chen et al., 2018; Roth et al., 2020). Draper et al. (2018) suggest using AI to supplement single operators by assisting in monitoring an autonomous vehicle's progress. However, increasing information per se will not automatically improve the operator's SA, information should help form a mental representation of the on-going process using techniques that integrate multi-levels of complexity (see Section 3.4 Multimodal Interfaces; Calhoun et al., 2018; J. Cha et al., 2019; Gillan et al., 2010). Also, SA becomes even more complex as agents become autonomous. Projection needs to become teleological—that is, having SA of the agent in relation to end-states and SA of the uncertainties involved in agents achieving the desired end state (Chen et al., 2018; Endsley 1995; Schaeffer et al., 2017).

4.2 Psychomotor Aspects

Remote navigation and remote manipulation (e.g., grasping, pushing, and payload management) are two key aspects of robot teleoperations (Steinfeld et al., 2006). For remote navigation, the operators often need to estimate the absolute sizes of objects or terrain characteristics so that they can decide whether it is safe for the robot to maneuver in the remote environment (e.g., without getting stuck in a depression) (Kanduri et al., 2005; Tittle et al., 2002). Human operators' spatial orientation and object identification in the remote environment also tend to be degraded (Darken et al., 2001). For example, studies on rescue robots (e.g., robots for search and rescue at the site of the World Trade Center after September 11, 2001) showed that human operators' performance was often compromised because of poor spatial awareness caused by inadequate video image from the cameras and/or sensors on the robots (Murphy, 2004). In real-world operations, operator performance is sometimes degraded even further due to robotic system failures. Among the challenges associated with teleoperations, communications and connectivity-related issues (e.g., limited radio signal strengths; Parasuraman et al., 2017) pose a special challenge for high stake environments (e.g., military missions and disaster relief) and can exacerbate human performance issues associated with teleoperation.

Latency refers to the delay between input action and (visible) output response, and is usually caused by transmitting information across a communications network (Allison et al., 2004; MacKenzie & Ware, 1993). Studies on human performance in VE show that people are generally able to detect latency as low as 10–20 msec (Ellis et al., 2004). Sheridan and Ferrell (1963) conducted one of the earliest experiments on the effects of time delay on teleoperating performance. They observed that time delay had a profound impact on the teleoperator's performance and the resulting movement time increases were well in excess of the amount of delay. A recent study shows that latency of 800 msec significantly degraded human operators' teleoperation task performance (Lu et al., 2019), and it was observed that, when system latency is over about 1s, operators tend to switch their control strategy to a "move and wait" one, instead of continuously commanding and trying to compensate for the delay (Lane et al., 2002). Based on this and other experimental results, Sheridan (2002) recommended that supervisory control and predictor displays be used to ameliorate the negative impact of time delays on teleoperation. Individual differences such as temporal sensitivity have been found to be predictive of teleoperation performance (Scholcover & Gillan, 2018). Furthermore, latency has been associated with motion/simulator sickness, which can be caused by cue conflict (i.e., discrepancy between visual and vestibular systems) (Kolasinski, 1995; Stanney, Mourant, & Kennedy, 1998). For example, in Oving and Van Erp's (2001) study on indirect driving of an armored vehicle, several participants in the HMD-driving condition (with a time delay) had to withdraw from the experiment due to motion sickness. Studies have also shown that high latency lags tend to reduce perceived telepresence (Ellis et al. 1999; Kaber et al., 2000). Compensation aids have been developed to assist the human operator's performance and workload when dealing with teleoperation with latency (Lu et al., 2019).

4.3 Workload

Given the multitasking requirements of HRI tasks and other perceptual/psychomotor-related challenges discussed above, the operator's workload is an important issue that needs to be considered. HRI tasks involving teleoperation can cause significant workload, particularly if there are suboptimal tasking environments, such as latency issues or loss of communications (Lu et al., 2019). For example, Lu et al. (2019) found that a time delay of 800 msec significantly increased human operators' workload in a teleoperation tasking environment. Even when working with autonomous robots and HRI tasks are highly automated, high workload can still result due to the cognitive demands associated with sustained focused attention required for monitoring and vigilance (Warm, Dember, & Hancock, 1996; Wohleber et al., 2019). Research has examined the roles of trait and state individual differences factors (Szalma & Taylor, 2011; Wohleber et al., 2019) and the relationship between distress and attention in HRI task performance (Matthews et al., 2017).

The workload issue is exacerbated in multirobot control situations where the operator needs to switch attention/control among the robots from time to time (Squire & Parasuraman 2010). Task switching costs can be problematic since the information processing required for dealing with one robot may interfere with dealing with another robot, either due to the similarity in task stimuli (e.g., video streams from the robots) and/or similarity in required task responses (e.g., map-based tasking/planning and spatial processing) (Kiesel et al., 2010). Models such as the Strategic Task Overload Management (STOM) have been investigated and shown to be useful to predict operator attention allocation (Wickens, Gutzwiller, & Santamaria, 2015). More specifically, task switch decisions appear to be largely based on relative task attractiveness, which is influenced by the perceived interest, difficulty, priority, and salience of the task. Wickens et al. (2016) tested the STOM model in a HRI multitasking experiment and the results showed that the model caused 95% of the variance in visual attention allocation.

Ways to mitigate workload issues have been investigated, such as Roth et al.'s (2020) workload-adaptive cognitive agent that can adaptively support helicopter crew's mission planning and execution in multirobot management (see Figure 8). Other interface design solutions have also been suggested to deal with the cognitive costs associated with task switching and interruptions (Norman, 1986). Finally, transparent HRI interfaces should take into account the potential workload implications associated with the additional information that the operator needs to process. So far, studies have shown that the additional information (e.g., based on the SAT framework) does not seem to increase the operators' workload in multirobot management contexts (Mercado et al., 2016; Roth et al., 2020; Stowers et al., 2020), although there is some evidence of higher workload associated with transparency displays in an aviation environment (Helldin, 2014).

4.4 Trust

Lee and See (2004) define trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (p. 54) and identify an agent's *purpose*, *process*, and *performance* as the antecedents for operator trust development. Trust implies a propensity to act by either accepting or rejecting an agent's course of action that may lead to errors of either misuse or disuse of the agent (Parasuraman & Riley, 1997). Trust is based on a confluence of factors—dispositional, situational, or learned attitudes towards robot-agents (Hoff & Bashir, 2015; Schaefer et al., 2016). Two meta-analyses of HRI experiments found that the most important predictor of trust is the performance aspect of the robot (Hancock et al., 2011), whereas human-related factors had a moderate effect on trust (Schaefer, Chen, Szalma, & Hancock, 2016). More specifically, robot performance-related factors (e.g., reliability, false alarm rate) were found to be better predictors of trust development than attribute-related factors (e.g., robot personality, anthropomorphism). Empirical findings also suggest that other factors such as the system's level of transparency and observability (e.g., stated intent and other items of the SAT framework) available to the human operator (Chen et al., 2018; Lyons et al., in press) and the operator's overall task load impact the operator's reliance on the agent (Manzey, Reichenbach, & Onnasch, 2012).

The issues related to trust in autonomous systems span multiple levels and encompass individual, social, cultural, and ethical aspects. Human operators' attitudes toward and trust in machine agents can be influenced by their own personalities (e.g., introversion versus extroversion) (Merritt & Ilgen, 2008; Szalma & Taylor, 2011), affective factors (e.g., moods and emotions) (Merritt, 2011), and states of stress, such as fatigue (Neubauer et al., 2012; Reichenbach, Onnasch, & Manzey, 2011). A recent study (Chien et al., 2020) on human interaction with a planning agent to manage multiple drones also demonstrates compelling cultural differences (the US, Taiwan, and Turkey) in operator trust when interacting with either transparent or opaque agents. Finally, operators may elect a manual solution even though they know the agent is more capable because they feel the outcome is the responsibility of the operator (Beck, Dzindolet, & Pierce, 2007). Safeguarding measures have been proposed to prevent autonomous systems from committing unethical or dangerous acts (Arkin & Ulam, 2012), and regulations-related issues have also been discussed (Cummings & Britton, 2019).

In summary, trust is not a unitary construct. Although greatly influenced by agents' performance, trust is also shaped by the operator's current situation, individual and cultural differences, and agents' characteristics (especially its transparency). De Visser et al. (2018) propose a design framework specifically to foster operator trust in autonomous agents. Computation-driven trust modeling (Kim et al., 2020; Nam et al., 2020) is also a promising approach to designing human-agent systems with trust considerations. As intelligent systems become increasingly sophisticated and are capable of learning/evolving either based on their learning algorithms or access to information from other networks (e.g., cloud-based systems), it is imperative to examine the implications of these capabilities on operator trust in the systems (Lyons et al., 2018). Since predictability is a critical aspect of trust development and maintenance, agent behaviors that change over time because of learning or new inputs from another network may prevent operators from properly calibrating appropriate trust. Recent research programs such as DARPA's eXplainable AI (Gunning & Aha, 2019) represent efforts to deal with these challenges.

4.5 Individual Differences

Effects of individual differences on cognitive task performance and interaction with automation have been well documented in the literature (Chen, 2011; Manzey et al., 2012; Parasuraman & Jiang, 2012; Shaw et al., 2010; Szalma & Taylor, 2011). Szalma (2009) suggests that, for human-technology interaction designs (including their associated training requirements), factors such as human variations in their characteristics and abilities should be taken into account. In fact, it has been demonstrated that effects due to individual differences in cognitive abilities can sometimes be even greater than effects due to interface design manipulations (Rodes & Gugerty, 2012). While there are many individual differences factors that can potentially impact HRI and there is even evidence of cultural differences (Chien et al., 2020), this section will focus on operator's spatial ability, attentional control, and working memory capacity.

4.5.1 Spatial Ability

Spatial ability has been found to be a significant factor in certain visual display domains (Stanney & Salvendy, 1995), multitasking involving flight asset monitoring and management (Morgan et al., 2011), navigation (Rodes & Gugerty, 2012), remote perception (Fincannon, 2013), visual search tasks (Chen 2011), and robotics task performance (Chen, 2011; Lathan & Tracey, 2002; Levulis et al., 2018; Liu et al., 2013). The U.S. Air Force (Chappelle, McMillan, et al., 2010; Chappelle, Novy, et al., 2010) identified spatial ability as an important factor for effective UAV control tasks (e.g., piloting and sensor operations), based on subject matter experts' interviews. Particularly, according to Liu et al. (2013), the spatial *visualization* and spatial *orientation* factors of spatial ability are more relevant to teleoperation performance than other factors. In military-related HRI contexts, several studies have shown that individuals with higher spatial ability performed significantly better and exhibited more effective visual scanning and target detection performance during HRI-related multitasking (Chen, 2011; Levulis et al., 2018). The effects of spatial ability on telerobotics task performance have also been demonstrated in the space domain (Liu et al., 2013). Rodes and Gugerty (2012) conclude that even though effective user interface designs compensate for low spatial ability in some spatial (UAV navigation) tasks, optimal task performance required both an effective interface and high spatial ability.

4.5.2 Attentional Control

Due to the multitasking requirements for most HRI tasks (e.g., sensor manipulation, tracking, communication; see Section 4.4), individual differences factors such as attentional control have been investigated in numerous studies (Chen, 2011; Levulis et al., 2018). Attentional control is defined as one's ability to focus and shift attention in a flexible manner (Derryberry & Reed, 2002), and there is evidence that those with better attentional control can multitask more effectively (Feldman Barrett, Tugade, & Engle, 2004; Rubinstein, Meyer, & Evans, 2001; Schumacher et al., 2001). Empirical evidence shows that operators with lower attentional control tend to rely more heavily on automation, even when the reliability of those systems are low (Chen & Barnes, 2012b; Chen & Terrence, 2009; Thropp, 2006). According to a U.S. Air Force's survey of subject matter experts, attentional control is one of the most important abilities that affect an operator's performance of UAV control tasks (Chappelle, McMillan et al., 2010). The importance of effective attention allocation for HRI task performance has been demonstrated in multiple studies (Chen 2011; Crandall and Cummings 2007). A recent study (Levulis et al., 2018) found that participants with higher attentional control scores, although not performing their HRI tasks significantly better, reported significantly lower workload; this finding is consistent with those reported in Chen (2011).

4.5.3 Working Memory Capacity

Since HRI tasks tend to involve multiple tasks simultaneously, as discussed previously, working memory capacity is an important individual differences factor due to its association with executive control processes and visual attentional allocation effectiveness (Ahmed et al., 2014; Bleckley et al., 2003). In fact, it has been reported that working memory capacity may be a better predictor of multitasking performance than attention (Bühner, 2006; Hambrick et al., 2010), especially visuospatial working memory (Logie, Trawley, & Law, 2011). In a multirobot management study, working memory capacity was found to modulate the effects of task load and network message quality on operator performance (Ahmed et al., 2014). Similar to low attentional control individuals, those with lower working memory capacity benefit more from reliable automation but are more negatively impacted when automation is less reliable, compared with those with higher working memory capacity (Rovira et al., 2017). Also, similarly, those with low working memory capacity tend to trust the automated system more than those with higher working memory capacity (Rovira et al., 2017). However, there appears to be an age effect that interacts with working memory capacity in human-automation interaction performance (Pak et al., 2017). Taken these findings together, HRI interface designs should take into account the individual differences aspects in order to achieve effective overall HRI system performance.

4.6 Training

Various training systems have been developed for HRI skills acquisition (Bric et al., 2016; Liu et al., 2013). Simulation-based training has been widely applied to provide trainees opportunities to practice HRI tasks in a safe and less costly manner. For example, in the medical domain, VR-based simulators are used to train robotic surgical skills using systems such as the da Vinci Surgery System (Bric et al., 2016). Bric et al. (2016) conducted a review of several commercially available VR trainers and found that VR-based simulation training was effective in supporting robotic surgical skills acquisition, although only limited data came from actual human patient cases. In the space domain, simulation-based training has also been investigated (Freer et al., 2020; Liu et al., 2013; Luko & Parush, 2017). Training is typically focused on those teleoperation tasks such as remote manipulation (e.g., robotic arm manipulation, docking of objects in desired locations), remote navigation (e.g., movement strategies to avoid collisions), and maintaining situation awareness (e.g., consolidating information from multiple cameras and frames of reference) (Liu et al., 2013).

Given the findings above on individual differences, it is also reasonable to expect that training interventions to deal with those individual differences may benefit HRI task performance, even if the training is not specifically developed for HRI applications. For example, training interventions that can help human operators better deal with their challenges in attention management and spatial reasoning have clear implications for HRI task performance (Sebok et al., 2019). There is also evidence that training can improve human operators' speed of information processing in multitasking situations (Dux et al., 2009). Additionally, to address multitasking requirements, a gaze-training paradigm reported in Wilson et al. (2011) to facilitate surgical skill acquisition could be beneficial for the HRI context. Training interventions that support attention management (e.g., visual scanning and prioritization, interruption management) and spatial cognition-related skill developments (e.g., for routing and tracking task performance) could also be fruitful (Baldwin & Reagan, 2009; Sebok et al., 2019; Spence & Feng, 2010).

There is extensive research on the connections between video gaming expertise and HRI (particularly unmanned aerial vehicle [UAV]) task performance. A U.S. Air Force study (McKinley, McIntire, & Funke, 2011) found that frequent video gamers outperformed infrequent gamers on robotics (UAV) tasks and, in some cases, performed as well as experienced pilots. Chen and Barnes (2012a, 2012b) also found that frequent gamers exhibited significantly better multitasking performance and SA of the tasking environments than did infrequent gamers. These findings are consistent with other studies that have shown video gamers' superior performance on tasks requiring visuospatial selective attention, multiple object tracking, rapid processing of visual information and imagery, flexibility in attention allocation, and other multitasking situations (Green & Bavelier, 2006; Hambrick et al., 2010). According to Spence and Feng (2010), playing action games benefits players' sensory, perceptual, and attentional abilities (e.g., contrast sensitivity, spatial resolution, attentional visual field, enumeration, multi-object tracking, and visuomotor coordination and speed), and this benefit is not only long-lasting but also transferable to tasks not closely similar to the tasks in the games (i.e., "far transfer"). The findings of the U.S. Military studies (Chen & Barnes, 2012a, 2012b; McKinley et al., 2011) support this hypothesis. Due to the implications of gaming for acquiring skills in visual attention allocation and visuospatial information processing, it is reasonable to consider the gaming aspects as either an operator selection consideration or a training tool. However, it is worth noting that Cummings et al. (2010) and Clare (2013) observed that frequent gamers exhibited different collaborative behaviors with intelligent agents (e.g., UAV planner) and tended to trust (or over-trust on occasions) the agents more than did infrequent gamers. Clare (2013) showed that transparency techniques (e.g., priming about agent performance/reliability) were effective in helping gamers more properly calibrate their trust in the planning agent when performing their multirobot control tasks.

5 CONCLUSION

The overall theme of the chapter is that role of the human in robotic operations is changing as robots become more intelligent and sophisticated. The emphasis of HRI has changed from considering robots as tools to considering robots as teammates that collaborate with humans in order to take advantage of the particular skill sets of each (not to mention the ethical and legal requirements of humans). However, manual control and its perceptual and cognitive requirements will remain important for the near future for safety and practical considerations. Capabilities to support effective human-robot teaming have progressed tremendously due to the advances in AI. This chapter reviews state-of-the-art HRI technologies and capabilities and key findings in HRI research, on which new advances in HRI can be based. Interdisciplinary research is advancing HRI in areas such as agent transparency/explainable AI, natural language-based and other intuitive interfaces, theory of mind, and shared mental models. Human factors scientists, working with researchers from other disciplines (AI, robotics, neuroscience, social sciences), will push the boundary of HRI to new frontiers such as ad-hoc human-robot teaming, human-robot joint training, and long-term human-robot symbiosis.

REFERENCES

Abiri, A., Pensa, J., Tao, A. et al. (2019). Multi-modal haptic feedback for grip force reduction in robotic surgery. *Scientific Reports, 9*, 5016.

Admoni, H., & Scassellati, B. (2017). Social eye gaze in human-robot interaction: A review. *Journal of Human-Robot Interaction, 6*(1), 25–63.

Ahmed, N., de Visser, E., Shaw, T. et al. (2014). Statistical modelling of networked human-automation performance using working memory capacity. *Ergonomics, 57,* 295–318.

Allison, R. S., Zacher, J. Z., Wang, D. et al. (2004). Effects of network delay on a collaborative motor task with telehaptic and televisual feedback. Association for Computing Machinery's (ACM) Special Interest Group on Computer Graphics and Interactive Techniques *(SIGGRAPH) Virtual-Reality Continuum and its Applications in Industry (VRCAI)* (pp. 375–381).

Andre, A. D., & Wickens, C. D. (1992). Compatibility and consistency in display control design: Implications for aircraft design. *Human Factors, 34,* 639–652.

Arkin, R., & Ulam, P. (2012). Overriding ethical constraints in lethal autonomous systems (Technical Report GIT-MRL-12-01). Atlanta, GA: Georgia Institute of Technology:

Baldwin, C. L., & Reagan, I. (2009). Individual differences in route-learning strategy and associated working memory resources. *Human Factors, 51*, 368–377.

Barber, D., Reinerman-Jones, L., & Matthews, G. (2015). Toward a tactile language for human robot interaction: Two studies of tacton learning and performance. *Human Factors, 57,* 471–490.

Barber, D., Wohleber, R., Parchment, A. et al. (2014). *Development of a squad level vocabulary for human–robot interaction*. Bielefeld, Germany: Springer-Verlag.

Barnes, M. J., Chen, J. Y. C., & Hill, S. G. (2017). Humans and autonomy: Implications for shared decision making in military operations. (Technical Report ARL-TR-7919). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Barnes, M. J., Chen, J. Y. C., & Jentsch, F. (2015). Designing for mixed-initiative interactions between humans and autonomous systems in complex environments. In *Proceedings of 2015 IEEE Conference on Systems, Man, and Cybernetics*, Hong Kong (pp. 1386–1390).

Barnes, M., Elliott, L., Wright, J. et al. (2019). Human robot interaction design research: From teleoperations to human-agent teaming (Technical Rep. ARL-TR-8770). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Barnes, M. J., Lakhmani, S., Holder, E., & Chen, J. Y. C. (2019). Issues in human–agent communication (Technical Report ARL-TR-8636). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Barnes, M. J., Wang, N., Pynadath, D.V. et al. (in press). Human-agent bidirectional transparency. In C. Nam & J. Lyons (Eds.), *Trust in human-robot interaction: Research and applications* Amsterdam: Elsevier.

Beck, H., Dzindolet, M., & Pierce, L. (2007). Automation usage decisions: Controlling intent and appraisal errors in a target detection task. *Human Factors, 49*, 429–437.

Bhaskara, A., Skinner, M., & Loft, S. (2020). Agent transparency: A review of current theory and evidence. *IEEE Transactions on Human-Machine Systems, 50*(3), 215–224.

Bleckley, M., Durso, F., Crutchfield, J. et al. (2003). Individual differences in working memory capacity predict visual attention allocation. *Psychonomic Bulletin & Review, 10*, 884–889.

Bodenhammer, A. (2007). Assessment of stereoscopic displays for assisting in route clearing operations. (Technical Report ARL-TR-4195). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Bohus, D., & Horvitz, E. (2019). Situated interaction. In S. Oviatt, B. Schuller, & P. Cohen (Eds.), *The handbook of multimodal-multisensor interfaces: Language processing, software, commercialization, and emerging directions* (Vol. 3, pp. 105–143). London: ACM.

Bonaiuto, S., Cannavò, A., Piumatti, G. et al. (2017). Tele-operation of robot teams: A comparison of gamepad-, mobile device and hand tracking-based user interfaces. *Proceedings of IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)* (pp. 555–560).

Breazeal, C. L. (2002). *Designing sociable robots*. Cambridge, MA: MIT Press.

Breazeal, C. L., Ostrowski, A., Singh, N. et al. (2019). Designing social robots for older adults. *The Bridge, 49*(1), 22–31.

Bric, J. D., Lumbard, D. C., Frelich, M. J. et al. (2016). Current state of virtual reality simulation in robotic surgery training: a review. *Surgical Endoscopy, 30*(6), 2169–2178.

Briggs, G., Williams, T., & Scheutz, M. (2017). Enabling robots to understand indirect speech acts in task-based interactions. *Journal of Human-Robot Interaction, 6*(1), 64–94.

Brown, D. S., Goodrich, M. A., Jung, S. et al. (2016). Two invariants of human-swarm interaction. *Journal of Human-Robot Interaction, 5*(1), 1–31.

Bühner, M., Konig, C., Pick, M. et al. (2006). Working memory dimensions as differential predictors of the speed and error aspect of multitasking performance. *Human Performance, 19*, 253–275.

Burke, J. L., Prewett, M., Gray, A. et al. (2006). Comparing the effects of visual-auditory and visual-tactile feedback on user performance: A meta-analysis. In *Proceedings of the International Conference on Multimodal Interfaces*. Alberta, Canada (pp. 108–117).

Calhoun, G. L., & Draper, M. (2010). Unmanned aerial vehicles: Enhancing video displays with synthetic vision technology. In M. J. Barnes & F. Jentsch (Eds.), *Human-robot interactions in future military operations* (pp. 229–249). Farnham: Ashgate.

Calhoun, G. L., Draper, M., Guilfoos, B. et al. (2005). Tactile and aural alerts in high auditory load UAV control environments. *Proceedings of Annual Meeting of Human Factors and Ergonomics Society, 49*(1), 145–149.

Calhoun, G. L., Ruff, H., Behymer, K. J. et al. (2018). Human–autonomy teaming interface design considerations for multi-unmanned vehicle control. *Theoretical Issues in Ergonomics Science, 19*(3), 321–352.

Catchpole, K., Bisantz, A., Hallbeck, M. S. et al. (2019). Human factors in robotic assisted surgery: Lessons from studies "in the Wild." *Applied Ergonomics, 78*, 270–276.

Cha, E., Kim, Y., Fong, T. et al. (2018). A survey of nonverbal signaling methods for non-humanoid robots. *Foundations and Trends in Robotics, 6*(4), 211–223.

Cha, J., Barnes, M. J., & Chen, J. Y. C. (2019). Visualization techniques for transparent human agent interface design. (Technical Report ARL-TR-8674). Aberdeen Proving Ground. MD: US Army Research Laboratory.

Cha, J. S., Monfared, S., Stefanidis, D. et al. (2020). Supporting surgical teams: Identifying needs and barriers for exoskeleton implementation in the operating room. *Human Factors*, *62*(3), 377–390.

Chakraborty, S., Tomsett, R., Raghavendra, R. et al. (2017). Interpretability of deep learning models: a survey of results. Paper presented at IEEE Smart World Congress 2017 Workshop: DAIS 2017—Workshop on Distributed Analytics Infrastructure and Algorithms for Multi-Organization Federations, San Francisco, CA (August 7–8, 2017).

Chandarana, M. (2019). Human-in-the-loop mission planning and monitoring for robot swarms. Doctoral dissertation. Carnegie Mellon University.

Chappelle, W., McMillan, K., Novy, P., & McDonald, K. (2010). Psychological profile of USAF unmanned aerial systems Predator & Reaper pilots. *Proceedings of Aerospace Medical Association of the 81st Annual Science Meeting*, Phoenix, AZ (May 9–13, 2010).

Chappelle, W., Novy, P., Randall, B., & McDonald, K. (2010). Critical psychological attributes of U.S. Air Force (USAF) Predator & Reaper sensor operators according to subject matter experts. In *Proceedings of the Aerospace Medical Association 81st Annual Science Meeting,* Phoenix, AZ (May 9–13, 2010).

Che, Y., Culbertson, H., Tang, C. et al. (2018). Facilitating human-mobile robot communication via haptic feedback and gesture teleoperation. *ACM Trans. Human-Robot Interaction, 7*(3), Article 20s.

Chen, J. Y. C. (2010). Robotic performance in a multi-tasking environment. In . M.J. Barnes & F. Jentsch (Eds.), *Human-robot* *interactions in future military operations* (pp. 293–314). Farnham: Ashgate.

Chen, J. Y. C. (2011). Individual differences in human-robot interaction in a military multitasking environment. *Journal of Cognitive Engineering and Decision Making, 5*(1), 83–105.

Chen, J. Y. C. & Barnes, M. J. (2012a). Supervisory control of multiple robots in dynamic tasking environments. *Ergonomics, 55*(9), 1043–1058.

Chen, J. Y. C. & Barnes, M. J. (2012b). Supervisory control of multiple robots: Effects of imperfect automation and individual differences. *Human Factors, 54*(2), 157–174.

Chen, J. Y. C. & Barnes, M. J. (2014). Human-agent teaming for multi-robot control: A review of human factors issues. *IEEE Transactions on Human-Machine Systems, 44*(1), 13–29.

Chen, J. Y. C., Barnes, M. J., & Harper-Sciarini, M. (2011). Supervisory control of multiple robots: Human performance issues and user interface design. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 41*(4), 435–454.

Chen J. Y. C., Haas, E. C., & Barnes, M. J. (2007). Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, 37*, 1231–1245.

Chen, J. Y. C., Lakhmani, S., Stowers, K. et al. (2018). Situation awareness-based agent transparency and human-autonomy teaming effectiveness. *Theoretical Issues in Ergonomics Science, 19*(3), 259–282.

Chen, J. Y. C., Oden, R. V. N., & Merritt, J. O. (2014). Utility of stereoscopic displays for indirect-vision driving and robot teleoperation. *Ergonomics,* 57(1), 12–22.

Chen, J. Y. C., Procci, K., Boyce, M., et al. (2014). Situation awareness-based agent transparency. (Technical Report ARL-TR-6905). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Chen, J. Y. C., & Terrence, P. I. (2009). Effects of imperfect automation and individual differences on concurrent performance of military and robotics tasks in a simulated multitasking environment. *Ergonomics, 52*(8), 907–920.

Chernova, S., & Thomaz, A. L. (2014). Robot learning from human teachers. *Synthesis Lectures on Artificial Intelligence and Machine Learning, 8*(3), 1–121.

Chien, S., Lewis, M., Sycara, K. et al. (2020). Influence of culture, transparency, trust, and degree of automation on automation use. *IEEE Transactions on Human-Machine Systems, 50*(3), 205-214.

Clare, A. S. (2013). Modeling real-time human-automation collaborative scheduling of unmanned vehicles. Doctoral dissertation. Massachusetts Institute of Technology.

Coa, L., Peng, C., & Hansberger, J. (2019). Usability and engagement study for a serious virtual reality game of lunar exploration missions. *Informatics, 6*(4), 44.

Cooke, N. J. (2015). Team cognition as interaction. *Current Directions in Psychological Science, 2*(6), 415–419.

Cooke, N., Demir, M., & Huang, L. (2020). A framework for human–autonomy teaming research. In *Proceedings of International Conference on Human–Computer Interaction* (HCII2020).

Couzin, I. D., Krause, J., James, R. et al. (2002). Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology, 218*, 1–11.

Crandall, J. W., Anderson, N., Ashcraft, C. et al. (2017). Human-swarm interaction as shared control: achieving flexible fault-tolerant systems. In *Proceedings of 14th International Conference on Engineering Psychology and Cognitive Ergonomics: Performance, Emotion and Situation Awareness,* Vancouver, Canada.

Crandall, J., & Cummings, M. L. (2007). Attention allocation efficiency in human-UV teams In: *Proceedings of American Institute of Aeronautics and Astronautics Infotech Aerospace Conference*.

Crandall, J. W., Goodrich, M. A., Olson, D. R. et al. (2005). Validating human–robot systems in multi-tasking environments. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 35*(4), 438–449.

Cummings, M. L., & Britton, D. (2019). Regulating safety-critical autonomous systems: past, present, and future perspectives. In R. Pak, E. J. de Visser, & E. Rovira (Eds.), *Living with robots: Emerging issues on the psychological and social implications of robotics* (pp. 119–141). Cambridge, MA: Academic Press.

Cummings, M. L., Clare, A., & Hart, C. (2010). The role of human-automation consensus in multiple unmanned vehicle scheduling. *Human Factors, 52*, 17–27.

Darken, R. P., Kempster, K., & Peterson, B. (2001). Effect of streaming video quality of service on spatial comprehension in a reconnaissance task. In *Proceedings of I/ITSEC 2001*, Orlando, FL (November 26–30, 2001) (pp. 391–400).

Defense Science Board (2016). Defense Science Board summer study on autonomy. Washington DC: Under Secretary of Defense.

de Looze, M. P., Bosch, T., Krause, F. et al. (2016). Exoskeletons for industrial application and their potential effects on physical work load. *Ergonomics, 59*, 671–681.

Derryberry, D., & Reed, M. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology, 111*, 225–236.

de Visser, E. J., Pak, R., & Shaw, T. H. (2018). From 'automation' to 'autonomy': The importance of trust repair in human–machine interaction. *Ergonomics, 61*, 1409–1427.

Dragan, A. D., Bauman, S., Forlizzi, J. et al. (2015). Effects of robot motion on human-robot collaboration. In *Proceedings of the 10th International Conference on Human-Robot Interaction (HRI'15)*, Portland, Oregon (March 2015). ACM (pp. 51–58).

Draper, M., Rowe, A, Douglass, S. et al. (2018). Realizing autonomy via intelligent adaptive hybrid control: adaptable autonomy for achieving UXV RSTA team decision superiority (also known as Intelligent Multi-UxV Planner with Adaptive Collaborative/Control Technologies *(IMPACT))* (Technical Report AFRL-RH-WP-TR-2018-0005). Wright-Patterson Air Force Base, OH: US Air Force Research Laboratory.

Dudek, G., Jenkin, M., & Milios, E. (2002). A taxonomy of multirobot systems. In T. Balch & L. E. Parker (Eds.), *Robot teams: From diversity to polymorphism* (pp. 3–22). Natick, MA: A. K. Peters.

Dux, P., Tombu, M., Harrison, S. et al. (2009). Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron, 63*, 127–138.

Dzindolet, M., Peterson, R., Pomranky, R. et al. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies, 58*(6), 697–718.

Edmondson, R., Light, K., Bodenhamer, A. et al. (2012). Enhanced operator perception through 3D vision and haptic feedback. In *Proceedings of SPIE, Unmanned Systems Technology XIV*, Baltimore, MD.

El-Shawa, S., Kraemer, N., & Sheikholeslami, S. (2017). "Is this the real life? Is this just fantasy?": Human proxemic preferences for recognizing robot gestures in physical reality and virtual reality. In *Proceedings of the Institute of Electrical and Electronics Engineers/Robotics Society of Japan International Conference on Intelligent Robots and Systems (IROS),* Vancouver, BC.

Elliott, L., Coovert, M.D., & Redden, E.S. (2009). Overview of meta-analyses investigating vibrotactile versus visual display options. In *Proceedings of the 13th International Conference on Human–Computer Interaction, Part II: Novel Interaction Methods and Techniques* San Diego, CA (pp. 435–443).

Elliott, L., Hill, S., & Barnes, M. (2016). Gesture-based controls for robots: overview and implications for use by soldiers. (Technical Report ARL-TR-7715). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Elliott, L., Jensen, C., Redden, E. S., & Pettit E. (2012). Robotic telepresence: perception, performance and user experience. (Technical Report ARL-TR-5928). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Elliott, L., & Redden, E. (2013). Reducing workload: a multisensory approach. In P. Savage-Knepshield (Ed.), *Designing soldier systems: Current issues in human factors*. Farnham: Ashgate.

Elliott, L., Skinner, A., Pettitt, R., Vice, J., & Walker, A. (2014). Utilizing glove-based gestures and a tactile vest display for covert communications and robot control. (Technical Report ARL-TR-6971). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Elliott, L., van Erp, J. B. F., Redden, E. et al. (2010). Field-based validation of a tactile navigation device. *IEEE Transactions on Haptics, 3*(2), 78–87.

Ellis, S. R., Adelstein, B. D., Baumeler, S. et al. (1999). Sensor spatial distortion, visual latency, and update rate effects on 3D tracking in virtual environments. In *Proceedings of Virtual Reality Conference* (pp. 218–221).

Ellis, S. R., Mania, K., Adelstein, B. D. et al. (2004). Generalizability of latency detection in a variety of virtual environments. In *Proceedings of the Human Factors Ergonomics Society 48th Annual Meeting* (pp. 2632–2636)

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

Endsley, M., & Kiris, E.O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*(2), 381–395.

Evans, A., Marge, M., Stump, E. et al (2017). The future of human robot teams in the Army: Factors affecting a model of human system dialogue towards greater team collaboration. In P. Savage-Knepshield & J. Chen (Eds.), *Proceedings of Advances in Human Factors in Robots and Unmanned Systems: Advances in Intelligent Systems and Computing, 499*. 197–209.

Everitt, T., & Hutter, M. (2018). Universal artificial intelligence. In M. Abbass & J. Scholz (Eds.), *Proceeding of foundations of trusted autonomy* (pp. 15–47). Cham: Springer Open.

Feldman Barrett L., Tugade, M., & Engle, R. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*, 553–573.

Fincannon, T. (2013). Visuo-spatial abilities in remote perception: A meta-analysis of empirical work. Doctoral dissertation. University of Central Florida.

Fitzgerald, T., Goel, A., & Thomaz, A. (2018). Human-guided object mapping for task transfer. *ACM Transactions on Human-Robot Interaction, 7*(2), Article 17.

Fong, T., Zumbado, J. R., Currie, N. et al. (2013). Space telerobotics: Unique challenges to human–robot collaboration in space. *Human Factors & Ergonomics Review – Human Performance in Teleoperations and Beyond, 9*(1), 6–56.

Fong, T., Thorpe, C., & Baur, C. (2003). Multi-robot remote driving with collaborative control. *IEEE Transactions on Industrial Electronics, 50*(4), 699–704.

Frasca, T., Oosterveld, B., Krause, E., & Scheutz, M. (2018). One-shot interaction learning from natural language instruction and demonstration. *Advances in Cognitive Systems, 6*, 159—176.

Freer, D., Guo, Y., Deligianni, F., & Yang, G. (2020). On-orbit operations simulator for workload measurement during telerobotic training. Online Journal arXiv, preprint. https://arxiv.org/abs/2002.10594

Fritsche, L., Unverzagt, F., Peters, J., & Calandra, R. (2015). First-person tele-operation of a humanoid robot. In *Proceedings of IEEE-RAS International Conference on Humanoid Robots* (*HUMANOIDS*) (pp. 997–1002).

Gartenberg, D., Breslow, L., McCurry, J. et al. (2014). Situation awareness recovery. *Human Factors, 56*(4), 710–727.

Gillan, D. J., Riley, J., & McDermott, P. (2010). The cognitive psychology of HRI. In M. Barnes & F. Jentsch (Eds), *Human-robot interactions in future military operations* (pp. 52–63). Farnham: Ashgate.

Giuliani, M., Szczęśniak-Stańczyk, D., Mirnig, N. et al. (2020). User-centred design and evaluation of a tele-operated echocardiography robot. *Health and Technology, 10*, 649–665.

Gombolay, M. C., Huang, C., & Shah, J.A. (2015). Coordination of human–robot teaming with human task preferences. In *AAAI Fall Symposium Series on AI-HRI*. Retrieved from https://interactive.mit.edu/sites/default/files/documents/Gombolay_2015_AAAI_FSS_AI-HRI.pdf

Goodrich, M. A., Crandall, J. W., & Barakova, E. (2013). Teleoperation and beyond for assistive humanoid robots. *Reviews of Human Factors and Ergonomics*, *9*, 175–226.

Goodrich, M. A., & Schultz, A. C. (2007). Human–robot interaction: A survey. *Foundations and Trends in Human Computer Interactions, 1*(3), 203–275.

Goodrich, M. A., Quigley, M., & Cosenzo, K. (2005). Task switching and multirobot teams. In L. Parker, F Schneider, & A. Schultz. (Eds.), *Multi-robot systems: From swarms to intelligent automata* (Vol. 3, pp. 185–195). Dordecht: Springer.

Green, C., & Bavelier, D. (2006). Enumeration versus multiple object tracking: The case of action video game players. *Cognition, 101*, 217–245.

Gromov, B., Abbate, G., Gambardella, L. M. et al. (2019). Proximity human-robot interaction using pointing gestures and a wrist-mounted IMU. In *Proceedings of 2019 International Conference on Robotics and Automation (ICRA)*, Montreal, QC, Canada (pp. 8084-8091).

Gunning, D. & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine, 40*, 44–58.

Haas, E. C., & van Erp, J. B. (2014). Multimodal warnings to enhance risk communication and safety. *Safety Science, 61*, 29–35.

Hall, E. T. (1966). *The hidden dimension*. New York: Doubleday.

Hambrick, D., Oswald, F., Darowski, E. et al. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology, 24*, 1149–1167.

Hancock, P. A., Billings, D. R., Schaefer, K. E. et al. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*(5), 517–527.

Hansberger, J. (2019). Real-time gesture recognition using 3D sensory data and a light convolutional neural network. In *Proceedings of 27th Association for Computing Machinery (ACM) International Conference on Multimedia* (October 21–25, 2019). Nice, France.

Hansberger, J., Peng, C., Blakely, V. et al. (2019). A multimodal interface for virtual information environments. In *Proceedings of 2019 Human Computer Interaction International Conference* (July 26–31, 2019). Orlando, FL.

Harris, J., & Barber, D. (2014). Speech and gesture interfaces for squad-level human-robot teaming. In *Proceedings of the Defense + Security Symposium*.

Hartnett, G., Elliott, L., Baraniecki, L. et al. (2018). Wearable concepts for gesture-based robot control, direction localization, and communication to a tactile belt display. (Technical Report ARL-TR-8597). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Helldin, T. (2014). Transparency for future semi-automated systems. Doctoral dissertation. Orebro University, Orebro, Sweden.

Hocraffer, A., & Nam, C.S. (2017). A meta-analysis of human-system interfaces in unmanned aerial vehicle (UAV) swarm management. *Applied Ergonomics, 58*, 66–80.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*, 407–434.

Holder, E. (2018). Defining soldier intent in soldier-natural language processing interaction. (Technical Report ARL-TR-8919). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Ibrahimović, R., Tsykunov, E., Shirokun, V. et al. (2019). DronePick: Object picking and delivery teleoperation with the drone controlled by a wearable tactile display. In *Proceedings of 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, New Delhi, India (October 14–18, 2019) (pp. 1–6).

Jiang, S., & Arkin, R. C. (2015). Mixed-initiative human-robot interaction: Definition, taxonomy, and survey. *IEEE International Conference on Systems, Man, and Cybernetics* (pp. 954–961).

Jung, M., & Hinds, P. (2018). Robots in the wild: A time for more robust theories of human-robot interaction. *ACM Transactions on Human-Robot Interaction, 7*(1), Article 2.

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.

Kaber, D. B., Riley, J. M., Zhou, R. et al. (2000). Effects of visual interface design, and control mode and latency on performance, telepresence and workload in a teleoperation task. In *Proceedings Human Factors Ergonomics Society 44th Annual Meeting* (pp. 503–506).

Kalaitzakis, M., Cain, B., Vitzilaios, N. et al. (2020). A marsupial robotic system for surveying and inspection of freshwater ecosystems. *Journal of Field Robotics,* 1–18.

Kanduri, A. K., Thomas, G., Cabrol, N. et al. (2005). The (in)accuracy of novice rover operators' perception of obstacle height from monoscopic images. *IEEE Transactions on Systems, Man, and Cybernetics, 35*(4), 505–512.

Karreman, D., Ludden, G., & Evers, V. (2019). Beyond R2D2: Designing multimodal interaction behavior for robot-specific morphology. *ACM Transactions on Human-Robot Interaction, 8*(3), 1–32.

Khan, A., & Anwar, Y. (2020). Robots in healthcare: A survey. In *Urban water management for future cities* (pp. 280–292). Cham: Springer.

Kiesel, A., Steinhauser, M., Wendt, M. et al. (2010). Control and interference in task switching: A review. *Psychological Bulletin, 136*(5), 849–874.

Kim, S., Moore, A., Srinivasan, D. et al. (2019). Potential of exoskeleton technologies to enhance safety, health, and performance in construction: Industry perspectives and future research directions. *IISE Transactions on Occupational Ergonomics and Human Factors, 7*(3–4), 185–191.

Kim, W., Kim, N., Lyons, J. B. et al. (2020). Factors affecting trust in high-vulnerability human-robot interaction contexts: A structural equation modelling approach. *Applied Ergonomics, 85*, 103056.

Kiribayashi, S., Yakushigawa, K., & Nagatani, K. (2018). Design and development of tether-powered multirotor micro unmanned aerial vehicle system for remote-controlled construction machine. In M. Hutter & R. Siegwart (Eds.), *Field and service robotics* (pp. 637–648). London: Springer.

Kolasinski, E. M. (1995). Simulator sickness in virtual environments. (Technical Report 1027). Alexandria, VA: US Army Research Institute.

Kolling, A., Walker, P., Chakraborty, N. et al. (2016). Human interaction with robot swarms: A survey. *IEEE Transactions on Human-Machine Systems, 46*(1), 9–26.

Kunze, A., Summerskill, S. J., Marshall, R. et al. (2019). Automation transparency: Implications of uncertainty communication for human-automation interaction and interfaces. *Ergonomics, 62*(3), 345–360.

Laaki, H., Miche Y., & Tammi, K. (2019). Prototyping a digital twin for real time remote control over mobile networks: application of remote surgery. *IEEE Access, 7*, 20325–20336.

Lane, J. C., Carignan, C. R., Sullivan, B. R. et al. (2002). Effects of time delay on telerobotic control of neutral buoyancy vehicles. *Proceedings of IEEE International Conference on Robotics and Automation, 3*, 2874–2879.

Lasota, P. A., Fong, T., & Shah, J. A. (2014). A survey of methods for safe human-robot interaction. *Foundations and Trends in Robotics, 5*(4), 261–349.

Lathan, C. E. & Tracey, M. (2002). The effects of operator spatial perception and sensory feedback on human–robot teleoperation performance. *Presence: Teleoperators Virtual Environments, 11*, 368–377.

Lee, J. D. (2012). Trust, trustworthiness, and trustability. Paper presented at the Workshop on Human–Machine Trust for Robust Autonomous Systems, Ocala, FL (January 31–February 2, 2012).

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80.

Levulis, S. J., DeLucia, P. R., & Kim, S. Y. (2018). Effects of touch, voice, and multimodal input, and task load on multiple-UAV monitoring performance during simulated manned-unmanned teaming in a military helicopter. *Human Factors, 60*(8), 1117–1122.

Lewis, M. (2013). Human interaction with multiple remote robots. *Human Factors & Ergonomics Review – Human Performance in Teleoperations and Beyond, 9*, 131–174.

Liu, A. M., Oman, C. M., Galvan, R. et al. (2013). Predicting space telerobotic operator performance from human spatial ability assessment. *Acta Astronautica, 92,* 38–47.

Logie, H., Trawley, S. & Law, A. (2011). Multitasking: Multiple, domain specific cognitive functions in a virtual environment. *Memory Cognition, 39*, 1561–1574.

Lu, S., Zhang, M., Ersal, T. et al. (2019). Workload management in teleoperation of unmanned ground vehicles: Effects of a delay compensation aid on human operators' workload and teleoperation performance. *International Journal of Human–Computer Interaction, 35*(19), 1820–1830.

Luko, L., & Parush, A. (2017). Targeted cognitive training of spatial skills: Perspective taking in robot teleoperation. In E. Clementini, M. Donnelly, M. Yuan, et al. (Eds.), *Proceedings of the 13th International Conference on Spatial Information Theory (COSIT 2017)*, 12: 12:1–12:13.

Lyons, J. B . (2013). Being transparent about transparency: A model for human–robot interaction. In D. Sofge, G. J. Kruijff, & W. F. Lawless (Eds.), *Trust and Autonomous Systems: Papers from the AAAI Spring Symposium* (pp. 48–53). Technical Report SS-13-07), Menlo Park, CA: AAAI Press.

Lyons, J. B. & Havig, P. R. (2014). Transparency in a human-machine context: interface approaches for fostering shared awareness/intent. In R. Shumaker & S. Lackey (Eds.), *Virtual, augmented and mixed reality: Designing and developing virtual and augmented environments* (pp. 181–190). Cham: Springer.

Lyons, J. B., Ho, N. T., Friedman, J. et al. (2018). Trust of learning systems: Considerations for code, algorithms, and affordances for learning. In F. Chen & J. Zhou (Eds.), *Human and machine learning: Visible, explainable, trustworthy and transparent* (pp. 265–278). London: Springer.

Lyons, J. B., Vo, T., Wynne, K.T. et al. (in press). Trusting autonomous robots: The role of reliability and stated social intent. *Human Factors*.

MacKenzie, S., & Ware, C. (1993). Lag as a determinant of human performance in interactive systems. In *Proceedings of Special Interest Group on Computer-Human Interactions (SIGCHI) Conference Human Factors Computer System* (pp. 488–493).

Mahey, C. E. V., Moses, L. J., & Pfiefer, J. H. (2014). How and why: theory of mind in the brain. *Developmental Cognitive Neuroscience, 9*, 68–81.

Manzey, D., Reichenbach, J., & Onnasch, L. (2012). Human performance consequences of automated decision aids: The impact of degree of automation and system experience. *Journal Cognitive Engineering Decision Making, 6*, 57–87.

Marquez, J., & Cummings, M. L. (2008). Design and evaluation of path planning decision support for planetary surface exploration. *Journal of Aerospace Computing Information Communication, 5*, 57–71.

Matthews, G., Lin, J., Panganiban, A. R. et al. (2020). Individual differences in trust in autonomous robots: Implications for transparency. *IEEE Transactions on Human-Machine Systems, 50*(3), 234–244.

Matthews, G., Reinerman-Jones, L., Abich, J. et al. (2017). Metrics for individual differences in EEG response to cognitive workload: Optimizing performance prediction. *Personality and Individual Differences, 118*, 22–28.

Mavridis, N. (2015). A review of verbal and non-verbal human–robot interactive communication. *Robotics Autonomous Systems, 63*(1), 22–35.

McKinley, R., McIntire, L., & Funke, M. (2011). Operator selection for unmanned aerial vehicles: Comparing video game players and pilots. *Aviation, Space Environmental Medicine, 82*, 635–642.

McNeese, N., Mustifa, D., Cooke, N. et al. (2017). Teaming with a synthetic teammate: insights into human-autonomy teaming. *Human Factors, 60*(2), 262–273.

Mercado, J., Rupp, M., Chen, J. et al. (2016). Intelligent agent transparency in human-agent teaming for multi-UxV management. *Human Factors, 58*(3), 401–415.

Merritt, S. (2011). Affective processes in human-automation interactions. *Human Factors, 53,* 356–370.

Merritt, S., & Ilgen, D. (2008). Not all trust is created equal: Dispositional and history-based trust in human–automation interaction. *Human Factors, 50,* 194–210.

Meszaros, E. L., Chandarana, M., Trujillo, A., & Allen, B.D. (2018). Compensating for limitations in speech-based natural language processing with multimodal interfaces in UAV operation. In J. Chen (Ed.), *Advances in human factors in robots and unmanned systems: Advancers in intelligent systems and computing 595,* (183–194). Cham: Springer.

Miller, C. A., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors, 49*(1), 57–75.

Morgan, B., et al. (2011). Individual differences in multitasking ability and adaptability. In *Proceedings of Human Factors Ergonomics Society Annual Meeting* (pp. 919–923).

Mortimer, B., Elliott, L. R., Miller, L. C. (2020). Designing tactile cues: Factors affecting perceived salience and recognition. *IEEE Transactions on Haptics*, *13*(4), 709–719.

Mueller, S., Hoffman, R., Clancey, W. et al. (2019) Explanation in human-AI systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for explainable AI. Report on Award No. FA8650-17-2-7711, DARPA XAI Program.

Mumm, J., & Mutlu, B. (2011). Human-robot proxemics: Physical and psychological distancing in human-robot interaction. In *Proceedings of ACM/IEEE International Conference on Human-Robot Interaction* (HRI), Lausanne (pp. 331–338)

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

Murphy, R. R. (2004). Human–robot interaction in rescue robotics. *IEEE Transactions on Systems, Man, and Cybernetics, Part C, Applications Review, 34*(2), 138–153.

Nam, C., Walker, P., Li, H. et al. (2020). Models of trust in human control of swarms with varied levels of autonomy. *IEEE Transactions on Human-Machine Systems, 50*(3), 194–204.

National Science Foundation. (2017). National robotics initiative 2.0: Ubiquitous collaborative robots. https://www.nsf.gov/pubs/2017/nsf17518/nsf17518.htm.

Neubauer, C., Matthews, G., Langheim, L. et al. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors, 54*, 734–746.

Norman, D. (1986). Cognitive engineering. In D. Norman & S. Draper (Eds.), *User centered system design: New perspectives on human-computer interaction* (pp. 31–61). Hillsdale, NJ: Erlbaum.

Obaid, M., Sandoval, E. B., Zlotowski, J. et al. (2016). Stop! That is close enough. How body postures influence human-robot proximity. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 354–361).

Olsen Jr, D. R., & Wood, S. B. (2004). Fan-out: measuring human control of multiple robots. In *Proceedings of the Special Interest Group on Computer-Human Interaction (SIGCHI) Conference Human Factors Computing System* (pp. 231–238).

Ophir-Arbelle, R., Oron-Gilad, T., Borowsky, A. et al. (2013). Is more information better? How dismounted soldiers use video feed from unmanned vehicles: Allocation and information extraction considerations. *Journal of Cognitive Engineering and Decision Making, 7*(1), 26–48.

Oron-Gilad, T. (2014). Scalable interfaces for operator control units: Common display to conduct MOUT operations with multiple video feeds. Ben Gurion University of the Negev (final research report). Be'er Sheva, Israel.

Oron-Gilad, T., & Parmet, Y. (2017). Close target reconnaissance: A field evaluation of dismounted soldiers utilizing video feed from an unmanned ground vehicle in patrol missions. *Journal of Cognitive Engineering and Decision Making, 11*(1), 63–80.

Oving, A. B., & Van Erp, J. B. F. (2001). Driving with a head-slaved camera system. In: *Proceedings of Human Factors Ergonomics Society Annual Meeting* (pp. 1372–1376).

Pacchierotti, C., Sinclair, S., Solazzi, M. et al. (2017). Wearable haptic systems for the fingertip and the hand: Taxonomy, review, and perspectives. *IEEE Transactions on Haptics, 10*(4), 580–600.

Pak, R., McLaughlin, A. C., Leidheiser, W. et al. (2017). The effect of individual differences in working memory in older adults on performance with different degrees of automated technology. *Ergonomics, 60*(4), 518–532.

Parasuraman, R., Barnes, M., & Cosenzo, K. (2007). Adaptive automation for human robot teaming in future command and control systems. *International Journal of Command and Control, 1*(2), 43–68.

Parasuraman, R., Caccamo, S., Baberg, F. et al. (2017). A new UGV teleoperation interface for improved awareness of network connectivity and physical surroundings. *Journal of Human-Robot Interaction, 6*(3), 48–70.

Parasuraman, R., Cosenzo, K., & de Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology, 21*, 270–297.

Parasuraman, R., & Jiang, Y. (2012). Individual differences in cognition, affect, and performance: Behavioral neuroimaging, and molecular genetic approaches. *NeuroImage, 59*, 70–82.

Parasuraman, R., & Manzey, D.H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52*(3), 381–410.

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

Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics Part A Systems Humans, 30*(3), 286–295.

Perez-Grau, F. Ragel, R. Caballero, F. et al. (2017). Semiautonomous teleoperation of UAVs in search and rescue scenarios. In *Proceedings of the International Conference on Unmanned Aircraft Systems* (ICUAS'17) (pp. 1066–1074).

Pettit, R. A., Redden, E. S., Carstens, C. B., & Hooper, D. (2012). Scalability of robotic controllers: effects of progressive autonomy on intelligence, surveillance, and reconnaissance robotic tasks. (Technical Report ARL-TR-6182). Aberdeen Proving Ground, MD: US Army Research Laboratory.

Pitman, D., & Cummings, M. L. (2012). Collaborative exploration with a micro aerial vehicle: A novel interaction method for controlling a MAV with a hand-held device. *Advances in Human-Computer Interaction,* 2012: 18.

Pollard, K. A., Lukin, S. M., Marge, M. et al. (2018). How we talk with robots: Eliciting minimally-constrained speech to build natural language interfaces and capabilities. In *Proceedings of the Human Factors and Ergonomics*.

Pynadath, D. V., Barnes, M. J., Wang, N. et al. (2018). Transparency communication for machine learning in human-automation interaction. In J. Zhou & F. Chen (Eds.), *Human and machine learning*. Cham: Springer.

Pynadath, D. V., & Marsella, S. C. (2005). PsychSim: Modeling theory of mind with decision theoretic agents. In *Proceedings of* *the International Joint Conference on Artificial Intelligence* (pp. 1181–1186).

Pynadath, D. V., Wang, N., Rovira, E. et al. (2018). Clustering behavior to recognize subjective beliefs in human-agent teams. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems*.

Redden, E. S., Elliott, L. R., Pettitt, R. A. et al. (2009). A tactile option to reduce robot controller size. *Journal of Multimodal User Interfaces, 2*(3-–4), 205–216.

Reichenbach, J., Onnasch, L., & Manzey, D. (2011). Human performance consequences of automated decision aids in states of sleep loss. *Human Factors, 53*, 717–728.

Riek, L. D. (2012). The Wizard of Oz paradigm in human robot interaction: A systematic review. *Journal of Human–Robot Interaction, 1*(1), 119–136.

Rodes, W., & Gugerty, L. (2012). Effects of electronic map displays and individual differences in ability on navigation performance. *Human Factors, 43*, 589–599.

Rosen, E., Whitney, D., Phillips, E. et al. (2019). Communicating robot arm motion intent through mixed reality head-mounted displays. *The International Journal of Robotics Research, 38*(12–13), 1513–1526.

Rosenfeld, A., Agmon, N., Maksimov, O., & Kraus, S. (2017). Intelligent agent supporting human–multi-robot team collaboration. *Artificial Intelligence, 252*, 211–231.

Roth, G., Schulte, A., Schmitt, F. et al. (2020). Transparency for a workload-adaptive cognitive agent in a manned–unmanned teaming application. *IEEE Transactions on Human-Machine Systems, 50*(3), 225–233.

Roundtree, K. A., Cody, J. R., Leaf, J. et al. (2020). Human-collective visualization transparency (preprint https://arxiv.org/abs/2003.10681)

Roundtree, K. A., Goodrich, M. A., & Adams, J. A. (2019). Transparency: Transitioning from human–machine systems to human-swarm systems. *Journal of Cognitive Engineering and Decision Making, 13*(3), 171–195.

Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions on Systems, Man, & Cybernetics, 22*, 1296–1308.

Rovira, E., Pak, R., & McLaughlin, A. (2017). Effects of individual differences in working memory on performance and trust with various degrees of automation. *Theoretical Issues in Ergonomics Science*, *18*(6), 573–591.

Rubinstein, J., Meyer, D., & Evans, J. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception Performance, 27*, 763–797.

Salas, E., Shuffler, M. L., Thayer, A. L. et al. (2015). Understanding and improving teamwork in organizations: A scientifically based practical guide. In . E. Salas, R, Rico, & J. Passmore (Eds.), *The Wiley Blackwell handbook of the psychology of team working and collaborative processes*. New York: Wiley.

Sanders, B., Vincenzi, D., Holley, S., & Shen, Y. (2019). Traditional vs gesture based UAV control. In J. Chen (Ed.), *Advances in human factors in robots and unmanned systems, advancers in intelligent systems and computing* (pp. 15–23). Cham: Springer.

Sato, R., Kamezaki, M., Niuchi, S. et al. (2020). Cognitive untunneling multi-view system for teleoperators of heavy machines based on visual momentum and saliency. *Automation in Construction, 110*, 103047.

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

Schaefer, K. E., Hill, S. G., & Jentsch, F. (2018). Trust in human-autonomy teaming: aA review of trust research from the US Army Research Laboratory. Robotics Collaborative Technology Alliance. In J. Chen (Ed.), *Proceedings of Applied Human Factors and Ergonomics, AISC 784* (pp. 102–114). Cham: Springer.

Schaefer, K. E., Straub, E. R., Chen, J. Y. C. et al. (2017). Communicating intent to develop shared situation awareness and engender trust in human–agent teams. *Science Direct Cognitive Systems Research, 46*, 26–39.

Scheggi, S., Aggravi, M., & Prattichizzo, D. (2017). Cooperative navigation for mixed human-robot teams using haptic feedback. *IEEE Transactions on Human-Machine Systems, 47*(4), 462–473.

Scholcover, F., & Gillan, D.J. (2018). Using temporal sensitivity to predict performance under latency in teleoperation. *Human Factors, 60*(1), 80–91.

Schulz, T., Torresen, J., & Herstad, J. (2019). Animation techniques in human-robot interaction user studies: A systematic literature review. *ACM Transactions on Human-Robot Interaction, 8*(2), Article 12.

Schumacher, E., Seymour, T., Glass, J. et al. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychology Science, 12*, 101–108.

Sebok, A., Walsh, M., Wickens, C. et al. (2019). Development of attentional skills training for operators of unmanned aerial systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 2161–2165).

Selkowitz, A., Lakhmani, S., & Chen, J. Y.C. (2017). Using agent transparency to support situation awareness of the Autonomous Squad Member. *Cognitive Systems Research, 46*, 13–25.

Shaw, T., Matthews, G., Warm, J. et al. (2010). Individual differences in vigilance: Personality, ability and states of stress. *Journal of Research Personality, 44*, 297–308.

Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. Cambridge, MA: MIT Press.

Sheridan, T. B. (2002). *Humans and automation: System design and research issues*. Hoboken, NJ: Wiley.

Sheridan, T. B. (2016). Human–robot interaction: status and challenges. *Human Factors, 58*(4), 525–532.

Sheridan, T. B., & Ferrell, W.R. (1963). Remote manipulative control with transmission delay. *IEEE Transactions on Human Factors in Electronics*, *4*(1), 25–29.

Sivčev, S., Coleman, J., Omerdić, E. et al. (2018). Underwater manipulators: A review. *Ocean Engineering, 163*, 431–450.

Skubic, M., Perzanowski, D., Blisard, S. et al. (2004). Spatial language for human–robot dialogs. *IEEE Trans Systems Man Cybernetics Part C Applications and Reviews, 34*(2), 154–167.

Son, H. I., Franchi, A., Chuang, L. et al. (2013). Human-centered design and evaluation of haptic cueing for teleoperation of multiple mobile robots. *IEEE Transactions on Cybernetics, 43*(2), 597–609.

Spence, I., & Feng, J. (2010). Video games and spatial cognition. *Review of General Psychology, 14*, 92–104.

Squire, P., & Parasuraman, R. (2010). Effects of automation and task load on task switching during human supervision of multiple semi-autonomous robots in a dynamic environment. *Ergonomics, 53,* 951–961.

Stankiewicz, P. G., Jenkins, S., Mullins, G. E. et al. (2018). A motion planning approach for marsupial robotic systems. In *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 1–9).

Stanney, K., Mourant, R. R., & Kennedy, R.S. (1998). Human factors issues in virtual environments: A review of the literature. *Presence: Teleoperators & Virtual Environments, 7,* 327–351.

Stanney, K. & Salvendy, G. (1995). Information visualization: Assisting low spatial individuals with information access tasks through the use of visual mediators. *Ergonomics, 38*, 1184–1198.

St Clair, A., & Mataric, M. (2015). How robot verbal feedback can improve team performance in human-robot task collaborations. In *Proceedings of International Conference on Human–Robot Interaction* (pp. 213–220).

Steinfeld, A., Fong, T., Kaber, D. et al. (2006). Common metrics for human-robot interaction. In *Proceedings of 2006 ACM Conference on Human-Robot Interaction* (pp. 33–40).

Stowers, K., Kasdaglis, N., Rupp, M.A. et al. (2020). The IMPACT of agent transparency on human performance. *IEEE Transactions on Human-Machine Systems, 50*, 245–253.

Szafir D. (2019). Mediating human-robot interactions with virtual, augmented, and mixed reality. In J. Chen & G. Fragomeni (Eds.), *Virtual, augmented and mixed reality: Applications and case studies. HCII 2019. Lecture Notes in Computer Science,* Vol. 11575. Cham: Springer.

Szafir, D., Mutlu, B., & Fong, T. (2017). Designing planning and control interfaces to support user collaboration with flying robots. *The International Journal of Robotics Research, 36*(5–7), 514–542.

Szalma, J. (2009). Individual differences in human-technology interaction: Incorporating variation in human characteristics into human factors and ergonomics research and design. *Theoretical Issues in Ergonomics Science, 10*, 381–397.

Szalma, J., & Taylor, G. (2011). Individual differences in response to automation: The five factor model of personality. *Journal of Experimental Psychology Applied, 17*, 71–96.

Tellex, S., Kollar, T., Dickerson, S. et al. (2011). Understanding natural language commands for robotic navigation and mobile manipulation. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*. San Francisco, CA (August 7–11, 2011). Palo Alto, CA: AAAI Press (pp. 1507–1514).

Thropp, J. E. (2006). Individual differences using automation. Doctoral dissertation. University of Central Florida.

Tittle, J. S., Roesler, A., & Woods, D.D. (2002). The remote perception problem. In *Proceedings of Human Factors Ergonomics Society 46th Annual Meeting* (pp. 260–264).

Valner, R., Kruusamäe, K., & Pryor, M. (2018). TeMoto: Intuitive multi-range telerobotic system with natural gestural and verbal instruction interface. *Robotics, 7*(1), 9.

Van Erp, J. B. F., & Padmos, P. (2003). Image parameters for driving with indirect viewing systems. *Ergonomics, 46*, 1471–1499.

Vasconez, J. P., Kantor, G. A., & Auat Cheein, F. A. (2019). Human–robot interaction in agriculture: A survey and current challenges. *Biosystems Engineering, 179*, 35–48.

Vause, N., Abouchacra, K., Letowski, T. et al. (2000). Spatial (3-D) audio displays improve speech communication in the command and control vehicle (C2V). In *Proceedings of the 22nd Army Science Conference*, Orlando, FL (pp. 834–841).

Veloso, M. (2018). The increasingly fascinating opportunity for human-robot-AI interaction: The cobot mobile service robots. *ACM Transactions on Human-Robot Interaction, 7*(1), Article 5.

Vered, M., Howe, P., Miller, T. et al. (2020). Demand-driven transparency for monitoring intelligent agents. *IEEE Transactions on Human-Machine Systems, 50*(3), 264–275.

Voshell, M., Woods, D. D., & Phillips, F. (2005). Overcoming the keyhole in human–robot coordination: Simulation and evaluation. In *Proceedings of Human Factors and Ergonomics Society Annual Meeting* (pp. 442–446).

Walker, M. E., Hedayati, H., & Szafir, D. (2019). Robot teleoperation with augmented reality virtual surrogates. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 202–210).

Wang, H., Lewis, M., Velagapudi, P. et al. (2009). How search and its subtasks scale in N robots. In *Proceedings of the ACM/IEEE International Conference on Human–Robot Interaction* (pp. 141–148).

Wang, N., Pynadath, D., & Hill, S. (2016). Trust calibration within a human–robot team: comparing automatically generated explanations. In *Proceedings of the 11th ACM/IEEE International Conference on Human Robot Interaction*. New Zealand, March 7–10. 2016, (pp. 109–116).

Wang, N., Pynadath, D., Rovira, E. et al. (2018). Is it my looks? Or something I said? The impact of explanations, embodiment, and expectations on trust and performance in human–robot teams. In *International Conference on Persuasive Technology* (pp. 56–69).

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds), *Automation and human performance: theory and applications* (pp. 183–200). Mahwah, NJ: Erlbaum.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*, 449–454.

Wickens, C. D., Gutzwiller, R. S., & Santamaria, A. (2015). Discrete task switching in overload: A meta-analysis and a model. *International Journal of Human Computer Studies, 79,* 79–84.

Wickens, C. D., Gutzwiller, R. S., Vieane, A. et al. (2016). Time sharing between robotics and process control: validating a model of attention switching. *Human Factors, 58*(2), 322–343.

Wickens, C. D., Prinet, J., Hutchins, S. et al. (2011). Auditory-visual redundancy in vehicle control interruptions: Two meta-analyses. In *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (pp. 1155–1159).

Williams, T., Szafir, D., & Chakraborti, T. (2019). The reality-virtuality interaction cube: A framework for conceptualizing mixed-reality interaction design elements for HRI. Paper presented at ACM/IEEE International Conference on Human-Robot Interaction.

Wilson, M., Vine, S., Bright, E. et al. (2011). Gaze training enhances laparoscopic technical skill acquisition and multi-tasking performance: A randomized, controlled study. *Surgical Endoscopy, 25*, 3731–3739.

Wohleber, R. W., Matthews, G., Lin, J. et al. (2019). Vigilance and automation dependence in operation of multiple unmanned aerial systems (UAS), A simulation study. *Human Factors, 61*(3), 488–505.

Woods, D. D., Tittle, J., Feil, M. et al. (2004). Envisioning human-robot coordination in future operations. *IEEE Transactions on Systems, Man, Cybernetics, Part C, Applications and Reviews, 34*(2), 210–218.

Wright, J. L., Chen, J. Y. C., & Barnes, M. J. (2018). Human–automation interaction for multiple robot control: the effect of varying automation assistance and individual differences on operator performance. *Ergonomics, 61*(8), 1033–1045.

Wright, J. L., Chen, J. Y. C., & Lakhmani, S. G. (2020). Agent transparency and reliability in human-robot interaction: The influence on user confidence and perceived reliability. *IEEE Transactions on Human-Machine Systems, 50*(3), 254–263.

Wu, C., Cha, J., Sulek, J. et al. (2020). Eye-tracking metrics predict perceived workload in robotic surgical skills training. *Human Factors, 62*(8), 1365–1386.

Wynne, K. T., & Lyons, J.B. (2018). An integrative model of autonomous teammate-likeness. *Theoretical Issues in Ergonomics Science, 19*, 353–374.

Xiao, X., Dufek, J., & Murphy, R. R. (2020). Tethered aerial visual assistance. *Journal of Field Robotics.*

Yanco, H. A., & Drury, J. (2004). Classifying human-robot interaction: an updated taxonomy. In *Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetics* (IEEE Cat. No.04CH37583: 3) (pp. 2841–2846).

Yena, J., Fana, S., Suna, S. et al. (2006). Agents with shared mental models for enhancing team decision makings. *Decision Support Systems, 41*, 634– 653.

Yogeeswaran, K., Złotowski, J., Livingstone, M. et al. (2016). The interactive effects of robot anthropomorphism and robot ability on perceived threat and support for robotics research. *Journal of Human-Robot Interaction, 5(*2), 29–47.

Young, S., & Peschel, J. (2020). Review of human–machine interfaces for small unmanned systems with robotic manipulators. *IEEE Transactions on Human-Machine Systems*, *50(*2), 131–143.

Zhao, J., Allison, R.S., Vinnikov, M., & Jennings, S. (2018). The effects of visual and control latency on piloting a quadcopter using a head-mounted display. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics* (SMC) (pp. 2972–2979).

Złotowski, J.A., Sumioka, H., Nishio, S. et al. (2018). Persistence of the uncanny valley. In H. Ishiguro & F. Dalla Libera (Eds.), *Geminoid studies* (pp. 163–187). Singapore: Springer Nature.