IEEE
SYSTEMS JOURNAL
A PUBLICATION OF THE IEEE SYSTEMS COUNCIL

Human Factors Considerations for Enabling Functional Use of Exosystems in Operational Environments

Leia Stirling[1,2], *Member, IEEE*, Ho Chit Siu[1], Eric Jones[3], Kevin Duda[3]
[1]Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Draper Laboratory, Cambridge, MA 02139, USA

The domain of human factors considers how to improve system design by considering the human within the design process, rather than designing a system and then considering the effects on the human after the design is completed. There have been decades of work in the domain of human factors and human-system integration in applications for complex systems to understand concepts within physical domains (e.g. system sizing and injury risk) and cognitive domain (e.g., workload, situation awareness, and automation mode confusion). Exosystem technologies are reaching the point where the transition to an operational environment is within reach. This paper presents human factors principles related to the cognitive domain. We present examples of these concepts in the context of exosystem design for operational environments and to aid the development of system evaluation standards. We also review current exosystem evaluation methods and present how performance metrics and tasks can be expanded to better describe system function in the context of known human factors complexities.

***Index Terms*—Exoskeleton, Human Factors, Wearable Computers, Human robot interaction, Performance Evaluation, Biomechanics, Training**

I. Introduction

HUMAN augmentation research is occurring in many different labs, with interest areas spanning mechanical design, material selection, and actuator and control system development. Dollar and Herr [1] present a review of lower extremity exoskeleton technology up to the time of its publication in 2008, with the years since seeing even greater expansion of work in wearable assistive systems [2], [3] (see Section III-A). Recently developed standards for personal care robots [4] include wearable robotics as a subset of the scope. The current standards provide hazard identification and risk estimation, particularly assessing stop features, workspace limitations, speed and force control, collision avoidance, and stability control. However, these standards do not provide a context for the underlying human factors concerns that may lead to risks and they do not articulate all the risks that are possible in complex integrated systems with a human in the loop. In the exosystem community, there is a need to define standards, as they provide a baseline set of design and performance criteria to compare different devices and to quantify the importance of design features with respect to operational application scenarios.

Using a morphological classification of robots [5], exosystems may be considered a class that is a "robot worn by a human to improve its performance or mitigate his handicap. This definition can be expanded to consider exoskeletons that may restore, enhance, or provide new human perceptual, cognitive, or physical abilities. Here we consider systems designed to assist motor performance, which are defined as enabling a user to reach an improved level of motor performance from their baseline. There are two aspects to motor performance: the physical action, including how the body moves and the amount of effort exerted or metabolic cost associated with the action, and the psychomotor mental workload imposed as mental resources are consumed to coordinate and perform the action [6], [7]. A device may enhance motor performance (increase what a user is capable of physically and mentally accomplishing at a given time) or restore performance (increase what a user is capable of accomplishing over time). In the latter case, the user may become independent from the device (such as with a rehabilitation system); while in the former case, the device would always be required to maintain augmented performance as the human is not capable of performing the task independently. In either of these modalities, the underlying powered components are controlled to modulate the motion of the user.

The intrinsic assumption is that the human adapts, or conforms, to work with the device to maximize performance, using biological muscles in conjunction with any external support or actuation. However, this may not always be the case. Gordon et al. [8] and Galle et al. [9] found that the initial response to their exoskeletons in able-bodied users was to fight the device, thereby increasing muscle activation, with the performance at the end of training showing reduced muscle recruitment for a sub-set of muscles. This adaptation period is variable, with some users having an easier time learning how to use a particular system. This example highlights the complexity of developing tightly coupled human-in-the-loop systems, where there is a time-varying response of the human to the system and the potential for different steady-state performance characteristics depending on the user.

The domain of human factors exists to improve system design by considering the human within the design process, rather than designing a system and then considering the effects on the human after the design is completed. Boff [10] considers four generations within the human factors domain. Generation 1 considers 'Physical Fit,' where equipment and the workplace are adapted to the human's capabilities.

Manuscript received XXX X, 2015; revised XXX. Corresponding author: L. Stirling (email: leia@mit.edu).
Digital Object Identifier 10.1109/JSYST.2018.2821689

Generation 2 considers 'Cognitive Fit,' which extends from the biomechanical requirements of Generation 1 to include challenges related to cognition. The shift in Generation 3 is to a 'Neural Fit,' which considers the fusion of technology with the human to augment human ability. The development of exosystems (e.g., exoskeletons and exosuits) falls within this framework (inclusive of considerations of Generations 1-3) as it is important to consider the biomechanical capabilities, cognitive constraints, and tightly-coupled interactions with technology to enhance and evaluate system designs. Considerations of the exosystem in this framework permit the evaluation of the mechanical and control systems at an integrated human-system level.

In this paper, we specifically consider the human factors concerns associated with cognitive fit in the framing of the exosystem domain. Although it is important to note that there is a coupling between these generations. Human factors for physical fit consider the anthropometry of the user population and the biomechanics of motion for the operational task. Common themes in biomechanics from a human factors perspective (also called occupational biomechanics) include understanding the kinematics and musculoskeletal loads for enhancing performance and minimizing injury [11]. These themes are already found in the exosystem literature, including evaluations of how systems affect kinematics, kinetics, muscle activations, and metabolics (discussed in Section III-A). Ferris et al. [12], provides a physiologist's perspective on exoskeleton design and discusses how kinematics, muscle activation, and energy consumption can be considered in the design of exosystems. Additional physical fit considerations relate to sizing and injury prevention, as well as thermal comfort.

Considerations for exosystems with regards to cognitive fit can inform the controller design architecture and evaluation methodologies. There has been significant work in the domain of human-centered automation considering the "man-machine system" [13] that should be used to inform the design of controls and displays for exoskeleton systems. A review of control system methods for exosystems [3] highlights that powered exosystems require a hierarchical control system, where different methods of control can be implemented at higher and lower levels of the controller. While these hierarchical models may implement a variety of control schemes, these models will require underlying modes (related to goals, tasks, or behaviors), where the different modes aid with different assist mechanisms (e.g., for lower extremity systems this could include controls that enable walking forwards, backwards, side-stepping on different terrains, or remaining stationary). While we use the term modes here, these modes may be reactive in nature, based on the measured state and driven by the estimated goal or task. For example, the use of sequential composition characterizes behaviors in the state space that are associated with relevant lower level controllers [14]. As the nature of the controller is behavior-dependent, we continue with the term mode.

Consider a lower extremity exosystem that can actively assist the knee during gait, but uses the knee angle to determine when to switch to a passive mode to enable ease with sitting. With this algorithm, the operator could begin a squat with active support, relying on the exosystem to assist the motion and yield a lower muscle activation. However, if he surpasses the threshold that triggers the passive mode, he may not be able to maintain the posture, losing balance and unintentionally falling. This example highlights the importance of having robust controllers along with human mode awareness and the need for understanding system modes with automated systems.

The trigger for switching between modes (which may also be considered as behaviors) can fall into four classes [15] (Table I). Exosystems and prosthetics in the literature have demonstrated controllers that change modes in each of these ways. For example, Varol et al. [16] developed a real-time supervisory control system for a lower-limb prostheses where mode changes were triggered by an intent recognition system trained on sensor data for when a person sits, stands, or walks (trigger class 3). The concept of a reactive controller, as in the case of sequential composition, would align with trigger class 3. It may be desired that an exosystem only operates within trigger class 3 and have seamless transitions between the underlying modes such that deliberate motions or stances are not required to obtain a particular behavior. The literature rarely discusses a systematic study of gracefully transitioning between task-based modes, instead focusing on performance within a particular task. Similar to what has been seen with other human-machine interactions [17], [18], there is potential for the user of an exosystem to become confused as to the current mode and thus behavior of the system or for the user to not remember how to achieve a specific behavior, highlighting the need to systematically study exosystem mode transitions.

While May et al. [19] found the adults are generally accepting of robots and are comfortable around robots, Jarrasse et al. [20] find that clinicians are regularly confronted with users who prefer mechanical cable-based devices or aesthetic limbs (i.e., systems that do not dynamically adjust) over more complex myoelectric active prosthetics. They postulate that this occurs since many socio-anthropological and cultural phenomena that may influence human-device interaction are not considered during the design and training process (e.g., the need for instant integration, the potential for new interaction modalities, and loss of versatility in task-specific devices). These particular concerns highlight the need for human factors considerations within the design process of assistive devices, including understanding the operational use case (e.g., is the system task specific or broadly usable).

TABLE I
TRIGGERS FOR SWITCHING BETWEEN MODES

Trigger Class	Description
1	The operator explicitly selects a new mode.
2	The operator enters data or a command that leads to a mode change (under all conditions, or when the automation, system, or environment is in a particular state).
3	The operator does not do anything, but transition is triggered by measured conditions according to algorithms built into the systems design
4	The operator selects a mode change, but automation does something else based on the underlying state, again as dictated by the systems control algorithms

The literature highlights many different fields that are now using concepts within human factors to improve human safety and performance. In the following sections, we present human factors principles that have the potential to guide exosystem design for use in an operational environment and to aid the development of system evaluation standards. Section II presents key concepts within the paradigm of cognitive fit and frames these examples in the context of exosystems. Section III presents a literature search and categorization of current exosystem evaluation methods and provides guidance on how performance metrics and tasks can be expanded to better describe system function in the context of known human factors complexities. While not explicitly evaluated in this paper, our motivating hypothesis is that exosystems that are designed following human factors principles and practices from the initial use-case definition will be more readily adopted and used in operational environments.

II. Application of Cognitive Fit Principles to Exosystems

A very generalized use case for exosystems would be to don the system, use the system, then doff the system. There is extensive literature in the domain of motor control that finds participant adaptation when exposed to altered force environments; and when the novel force environment is removed, errors are made as the participant readapts to the nominal environment (e.g., [21]–[23]). This adaption mechanism has been observed in exosystems [9], [24], [25] and has been explored as a way to aid motor learning during rehabilitation [24]. The after-effects for a given exosystem may occur within the usage of the system during transitions between system modes and during off-nominal conditions (i.e., conditions that are unexpected and undesired; see Section II-C). While some systems may have unintended after-effects, it may be the case that users develop a dual-adaption over time such that they can transition easily between wearing and not wearing the exosystem. The development of a dual-adaptation is seen in Gordon et al. [25], where use of the ankle exosystem on day 2 had a reduced adaptation timescale compared to day 1. Motor adaptation of posture, arm reaching, and locomotion have been extensively studied, including in altered force environments (e.g., [21], [22], [26]). While not a direct analog to exosystems, these studies highlight the ability of a person to adapt and transition to different loads on the body.

In this section, we provide additional background on cognitive fit considerations and specifically apply these constructs to the exosystem. For additional review of these concepts, there are several texts that go into additional detail for other application areas (e.g., [27]–[31]).

A. Mental Models

Mental models have been defined in many different ways. Here we use the definition from Rouse and Morris [32] that “mental models are mechanisms whereby humans generate descriptions of system purpose and form, explanations of system functioning and observed system states, and prediction of future system states.” Mental models are an evolving memory structure that provide a dynamic representation of the environment, as well as descriptive interrelationships for a set of objects or events. Our mental models can be developed through training and are used to inform our decision making through anticipation of system response and selection of behavior. For exosystem operations, the wearer needs know in what scenarios it can operate, what inputs they must provide, and what responses will be produced.

Rasmussen [33] considers operator performance at three levels–skill-based behavior (sensory-motor performance taking place without conscious control), rule-based behavior (stored procedures are implemented at a conscious level), and knowledge-based behavior (goals are known, but knowledge must be used to form a procedure). At the beginning of exosystem training when the system is not well-known, the operator would use knowledge-based or rule-based behavior to create the planned actions. However, over time the usage of the exosystem should take place without conscious control as a skill-based behavior. There is opportunity for defining strategy training that can directly provide the appropriate rules, and can be used to develop the mental model to enable improved procedure development in off-nominal scenarios. In contrast to training humans to use most other robotic systems where off-board demonstration provides visual feedback for a user to develop a mental model, the vast majority of user feedback in an exosystem is tactile rather than visual. This difference changes the requirements for mental model development in physically coupled systems such as these, as direct use of the system is required to learn the system.

Norman [34] observes that mental models are typically incomplete, can be unscientific, are unstable (forgetting occurs), and do not have firm boundaries (similar systems can cause confusion). This means that the user’s mental model does not necessarily match the conceptual device model. However, conflict between a user mental model and system model can result in poor operational performance. For example, the Human Universal Load Carrier (HULC) was designed to off-load soldier-borne loads and assist lower limb advancement during gait. The operators could not use natural kinematics, with observations that the system would try to advance the leg when the human operator was not ready [35]. This created a system that was operating differently than anticipated by the human, creating increased rather than decreased energy consumption during operation. The development of relevant feedback to the user, such as automation behavior (e.g., system status, goals, projected states) or expected system inputs, during training could help build this model [36] such that less feedback is required during operational use once expectations are developed.

Within a system, there are multiple embodied models that must be considered. In addition to the operators mental model of the system, there is the model used for the display to the user, the embedded models in the system software, and the model used for task control [13]. Inappropriate mappings between these underlying models can cause reductions in performance and increases in the frequency and magnitude of errors due to confusion, distraction, and concerns for safety [37], [38]. Zhang et al. [39] and Ding et al. [40] have shown that different control methods and timings of active assistance altered human performance. It may be that certain control mappings better align with the operator's mental model across goals or behaviors, yielding improved performance and the ability to develop improved skill-based behaviors.

B. Attention, Workload, and Situation Awareness

The concepts of attention, workload, and situation awareness (SA) are not tangible, but are mental constructs that aid in describing and understanding human cognitive performance. Attention is awareness directed at an element in the environment. The manner in which the focus of attention is selected and why that focus may shift are important considerations. In a research lab environment, where a user can operate an exosystem without external distraction, we may see differences in performance than when the exosystem is used in an operational environment where attention becomes divided. There are many models of attention that highlight how it is allocated and divided [29]. Several suggest there are thresholds that, when passed, lead to reductions in performance, and that optimizing multitask human-system interaction requires planning. Features from these models can aid in developing evaluation methods for exosystems.). If a researcher wants to evaluate how an exosystem affects central processing, it would be important to consider tasks that provide the relevant interference to assess these limitations. This is explored further in Section III-B.

The interference of tasks causes increased mental workload, which can be described as the level of attentional resources required for an operational task. This demand for resources can include the perception of visual, auditory, or haptic stimuli; cognitive decision making processes; and psychomotor coordination [6], [7]. The relationship between attention and performance is typically described as an inverted u-shape, with low and high attention states yielding low quality of performance, although the shape of the curve can change with complexity of the task [41]. With high workload, there is an increased tendency to focus only on a few relevant cues and the ability to voluntarily shift attention is reduced [42]. If an exosystem requires increased mental workload to operate, for example by requiring defined postures to trigger mode changes or to verify that automated mode changes were appropriately transitioned, additional attentional resources are directed towards the task of operating the system. This would have potential to limit and degrade any additional tasks that the user performs (e.g., interpreting information from the environment for a surveillance task or avoiding obstacles).

Design decisions and training methods can affect a users ability to interpret relevant information. Here we use Endlsey's decomposition of SA into three levels [43]: (1) perception of the elements, (2) comprehension of the current situation, and (3) projection of future status. While a user may be able to perceive information in the environment (Level 1), it may not be apparent how these cues would affect the use of the exoskeleton (Level 2 and 3). Thus, an inappropriate action may be taken. Breakdowns in SA can occur on any of the three levels and therefore evaluation of SA in the context of exosystem use is important for understanding the ability of a user to make operational decisions. Consider an exosystem ankle that nominally actively assists rotation based on the interaction force recorded with the ground, except in a separate mode where the interaction force is used to stiffen the joint and limit motion. If the user perceives (Level 1 SA) cues that lead to projecting (Level 3 SA) that the exosystem should be actively assisting, but the joint instead stiffens, the user would expect assistance and may then lose balance and fall when the joint stiffens instead.

As one may notice, there is an interplay between these constructs. A user's attention will be directed based on their underlying workload and their mental model of the system. However, the comprehension of the perceived environmental information and projection of that information will affect any ensuing decisions. Response from one's actions affects the state of the environment, which must then be perceived. SA can be used to revise and improve the mental model, but the mental model also helps to direct comprehension and projection. Maintaining relevant SA becomes difficult as the environment and/or system increase in complexity. Conflict can exist between perceived information and what a person thought should happen based on the mental model. Many current exosystems are designed to operate in a steady state environment. Transitioning to different tasks (e.g., walk to run or standing) or environments (e.g., level ground to inclines or stairs) may require a change in the underlying controller or actuator parameters. Even if the user does not need to manually change the system, the user does need to be aware of the effect of the change to prevent risks of loss of balance or injuries.

It is evident that an exosystem should require minimal user input, while maintaining appropriate user SA of the system, as this has implications in trusting the automation (discussed further in Section II-C). It has been shown that displayed information that is synergistic with a user mental model can improve performance, implying decreased workload to integrate the information [44]. However, SA and performance are not always correlated [43]. One may have high performance with high or low SA depending on the task. Thus a balance must be made in providing information that is relevant to task performance without overloading the user. It is also important to distinguish between SA of the system and SA of the environment when considering what is presented the user. Many exosystems do not have displays to provide the user with feedback, limiting the ability to provide system SA to the user. What could be helpful is to create a display for training purposes to enable the user to build an appropriate mental model such that a display is no longer needed during operational use. Alternatively, it may be observed that there are certain types of information that a user does need for operational use and that displays may be warranted (e.g., heads-up visual, a heads-down screen, or tactile feedback).

The ability for the exosystem to provide feedback to the human or adapt its response based on interrogating the humans intent has not been explored (Section III-A) and could be a way to improve human-exosystem fluency. In scenarios where there is a robot that is not physically in contact with the human, studies have specifically examined using feedback (e.g., through gestures [45]), verbal communication [46], non-verbal communication [47], and visual displays [48]) to enable the human to understand the robotic system. There is a distinction here between displays that may benefit training versus those that may benefit operations. During the training process the goal is to aid the user in adapting to efficient use of the system, i.e. enabling the development of the mental model and the appropriate motor patterns. A standardized training could include more information presented on actuation timings to permit the user to more easily entrain during steady state cyclic motions, or information could be provided on how the onboard sensing and underlying user biomechanics trigger changes in the underlying modes (see Section II-C for discussion of modes). In an operational environment, the additional information presented would be minimized to data relevant for efficient performance in the selected use case.

C. Automation, Mode Confusion, and System Trust

The control of exosystems has ranged from manual control (direct input from the user) to fully automatic control (operates without human input) [3], [37]. Sheridan [13] originally developed the concept of "supervisory control," where the human provides inputs to control some goals, while an embedded computer controls other goals (Fig. 1). Manual control is defined as the human directly controlling and receiving information from the system, with sensing or actuating directly transmitted, or with a computer transformation of the data. In these first two systems, all decisions are made by the human and the computer acts in an open-loop manner. In systems with supervisory control, there are decisions that are not made by the human and the computer acts in a closed-loop manner. The previously-mentioned prosthesis developed by Varol et al. [16] automatically closed lower-level control loops related to actuation, but the actions of the user (as detected by the system's sensors) determined the high-level mode of operation (defined previously as mode transition class 3 and is expressed as supervisory control in this framework). The human operator does not directly control the underlying actuator command, that behavior is determined via the embedded computer algorithms. The fully automatic control has the human as a pure observer, providing no influence on the system. Fully automatic systems have been explored in wearable technology through fixed-base rehabilitation exosystems for the lower extremity as well as in upper extremity systems, where the system moves the human without the human signaling the action, thus guiding the motion of the operator. However, as discussed in more detail by Marchal-Crespo and Reinkensmeyer [49], studies have found that engaging the patient in the control loop is preferred for regaining motor function.

The level of system automation (the degree to which the computer completes the task) does not need to be fixed. Parasuraman, Sheridan, and Wickens [50] simplify the human information processing model into four stages (sensory processing, perception/working memory, decision making, and response selection) and highlight that the level of automation even within a single system can vary based on the stage, permitting assistance in information acquisition, information

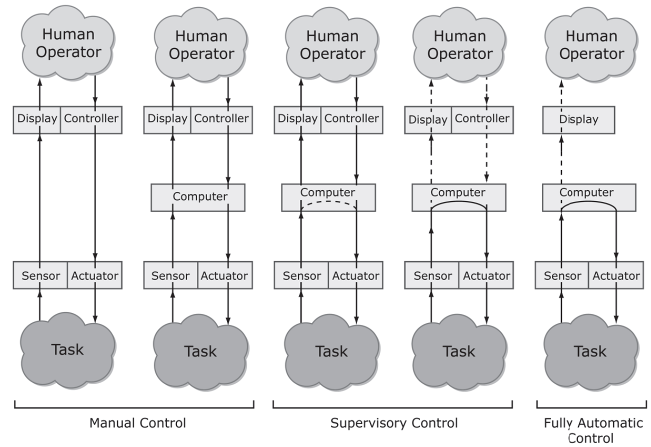

Fig. 1. Schematic of Sheridan's levels of automation. This is not a preferred ordering, or even a level of sophistication. These levels are ordered based on the types of interaction between the human and system. Dotted lines represent minor loop closures while solid lines are major loop closures. (Image: MIT OpenCourseWare)

analysis, decision and action selection, and action implementation. Further, the levels of automation within a stage can be adaptive, permitting a dynamic change in authority between the human and the system. For an integrated exosystem, there is the ability for the system to be automated at a high or low level within each of these stages. Table II highlights examples of the automation spectrum that are possible (based on the Hart and Valasek [51] levels of autonomy), although clearly the lowest level of automation would not be preferred for an active exosystem. For training purposes, there may be a desire for the system to be at a mid-level of automation that enables mental model development.

Research in human-robot teaming–where the robot and human are not physically connected–has explored how the degree of robot autonomy affects human SA ([52], [53]), particularly with regard to human decision making in cases of control reallocation. Scerri et al. [54] consider a methodology for making these transfer-of-control decisions while mitigating coordination failures within the team. While studies have found humans prefer greater autonomy (e.g., [55], [56]), there is a risk in decreasing relevant SA cues for the human. In teleoperated robotics, research is specifically looking at how to provide additional SA cues to the human [57]. However, these studies have typically occurred with a stationary human, where the supervision and actions do not require locomotion. Thus, there is a clear gap to understand how SA is affected during physical activity in which the robot is physically connected to the human, as in the case of an exosystem, and how reductions in human SA affect operational performance.

As operational tasks become more complicated and varied, the ways in which the human can interact with the system may become confusing. Mode awareness, the ability of the supervisor to track and anticipate system behavior, is quite important. Sarter [17] and Wood [18] provide summaries of the extensive research in automation for an aircraft flight deck, specifically the issue of mode awareness. As shown in Table

TABLE II
EXAMPLES OF AUTOMATION LEVEL FOR AN EXOSYSTEM

Processing Stage	Automation Level		
	Low	Mid	High
Information Acquisition	The user monitors all incoming sensor data or visual inputs to understand the state of the system	The exosystem displays relevant information for the human to monitor	The exosystem collects all data without displaying any to the user
Information Analysis	The user interprets all data	The exosystem analyzes the data and shows the user predictions of what may occur	The exosystem analyzes and interprets data and does not display to the user
Decision and Action Selection	The user ranks potential options	The exosystem ranks options and provides feedback on why a decision was made	The exosystem ranks and selects option without displaying the result to the user
Action Implementation	The human manually controls the degrees of freedom of the exosystem based on the decision made	The exosystem enacts the decision made, but gives the user a context-dependent time to veto	The exosystem executes the decision made without user interaction

I, changes in mode can be triggered in different ways [15]. Because a user has limited attentional resources, there needs to be a consideration of the user's environment and workload when determining how much direct user input is required. As previously mentioned, the less the user must directly input and control, the better for synergy and ability to engage with the surroundings. However, the selection of the underlying controller impacts the kinematics and metabolic consumption (e.g., [40], [58]–[60]). The actuation timings are important for enabling efficient mode transitions such that systems do not have a sluggish response or disturb the desired biomechanics. Thus, they also provide design requirements for control system development and highlight that beyond setting bounds on the controller lag times, there is a need for intent-based controllers to mitigate the lag inherent in feedback control and ensure the system responds according to the operators intentions. The need to enable intent-based control was also emphasized by Anam and Al-Jumaily [3]. The capability for anticipatory and feedforward control is an ongoing area of research. In human-robot teams, studies find that the ability to anticipate human behavior can improve the safety and fluency between a human and robot (e.g., [61]–[63]).

With less interaction from the user, there are risks associated with not having the appropriate SA of the current system mode. Sarter et al. [64] highlight that automation risks can originate from lack of mode awareness, gaps in mental models of how various automated modes work and interact, coordination errors in off-nominal situations, and overtrust in the automation. Leveson et al. [15] further decompose sources of mode confusion as highlighted in Table III. If we specifically consider the exosystem, we see these sources of mode confusion can appear in exosystem architectures as well (Table III). Limitations of an exoskeleton in either hardware or software, while undesirable, will exist in early systems. Considerations of the interactions between the hardware and software in the context of these known automation concerns can help drive the system design, training of the system, and create appropriate operator expectations.

While the putative benefits of automation are that it can free up resources, require less knowledge, and reduce human error, there are real complexities that arise as the automation creates new cognitive work [64]. It is therefore important to consider off-nominal conditions that may be encountered, in addition to nominal conditions, and how to ensure the system degrades gracefully when they do occur. In aviation, the "gracefulness" of a specific mode transition refers to the magnitude of undesirable decreases in operator performance, increases in workload, and changes in SA associated with that mode transition [65]. Graceful transitions are desired whether the transition is expected or unexpected, as they minimize the chances for an automation surprise.

The adoption of and reliance on wearable assistive technologies in an operational environment requires an appropriate trust in the automated system. Misuse, or overreliance on automation, can result in engaging in tasks for which the system was not designed, whereas disuse is the underutilization of the system and defeats the purpose of having an active assistive aid. The possibility of these inappropriate levels of trust means there needs to be a calibrated trust between the automation capability and the user belief [66]. Parasuraman and Riley [67], Lee and See [66], and Madhavan and Wiegmann [68] provide suggestions to designers in terms of both system design and user training so that systems can be made trustable and thus used. Exosystem users must be made aware of the limits of the technology, whether it is the specific set of tasks that the system is designed for, or the physical hardware or software safety limits that are designed into the system. Through increased operational use, and incremental demonstrated successful scenarios with the system, the user can increase their trust in the system. This trust can be built up by providing feedback to the user that is comprehensible in the context of the application and user goals, and includes training operators on expected governing behavior and intended use states. It can also be built up through demonstrated reliability and availability of the system for the required tasking. During training, revealing how context affects automation capability provides an improved mental model for the user, which Lee and See [66] term "trust resolution" for automation. If we consider a particular exosystem that is optimized for and only active in straight and level environments, and the user is only trained in that environment, the user might misuse the automation to navigate irregular terrain if he tries to

TABLE III
SOURCES OF MODE CONFUSION

Source*	Description*	Exosystem Examples
Interface interpretation errors	Computer interprets user-entered values differently than intended, or maps multiple conditions to the same output due to the same interface being used for different plant states.	Different errors are presented (e.g., visual, audio, or haptic display) using the same alarm message. For example, the alarm signifying the system estimates there is not enough torque to achieve an anticipated goal is the same as the indication that the threshold for safe extension has been reached for the estimated loading. The operator's action may differ depending on why he believes the system is presenting the alarm.
Inconsistent behavior	Behavior is inconsistent between modes.	An exosystem ankle DOF is nominally actively assisted to rotate based on the interaction force recorded with the ground, except in a separate mode where the interaction force is used to stiffen the joint and limit motion.
Indirect mode changes	Automation changes mode without an explicit instruction by the operator.	If the exoskeleton has multiple control assist strategies based on the estimated goal (e.g., different algorithms for stairs, steady-state gait, knee bends, or steady-state balance assist) the users may have a mental model of what the system will do in a particular scenario based on their intent. If the user changes posture unintentionally and that signal is mistaken for a mode change that is not made explicit, the user may not understand why the system changes behavior.
Operator authority limits	Interlocks and lockouts to ensure safety may affect off-nominal needs.	The exosystem may contain interlocks within the software or hardware to prevent joints from moving too quickly; however, in an off-nominal scenario such as avoiding an oncoming bike or car, the user may need to move quicker than permitted by the system.
Unintended side effects	Action intended to have a particular effect has an additional unanticipated effect.	If the exoskeleton is designed to have performance based on a particular alignment with anatomical landmarks, but becomes misaligned during use, the system may not appropriately interpret sensor readings and will respond incorrectly or not at all.
Lack of appropriate feedback	Inappropriate feedback on authority limits, system state, etc.	An exosystem may use more energy augmenting performance of certain task subsets when compared to others. If the user does not have a good mental model of the system behavior and energy usage, the user may select a poor plan of action if power consumption information is not presented in some manner.

* Summarized from Leveson et al. [15]

transfer the previous successful training experience. This error would be a result of having poor trust resolution, where actual appropriate trust in one mode leads to a perception of appropriate trust in another where the user actually has no prior experience.

III. Selecting Relevant Performance Metrics and Tasks

Prior to discussing how human factors concepts can guide an exosystem evaluation framework, we must first understand how exosystems are currently evaluated. In Section III-A, the current methods of measuring exosystem performance are reviewed. Section III-B follows with important considerations for designing and evaluating exosystems with a human factors framework. These sections emphasize the design of human studies to evaluate and characterize the human-exosystem interaction. In addition to these methods, there is also considerable need for improved human-exosystem modeling efforts to enable model-based designs and safety evaluations. While human studies can provide valuable information on a particular use-case, they do not permit a straightforward way to examine design decisions nor a full understanding of potential safety hazards [69]. Current modeling efforts lack the ability to predict the coupled kinematics of the human with an exosystem in parallel, where the exosystem can move independently from the human (i.e., permitting shearing at the human-machine interface) and the human kinematics are affected by the estimated interaction.

A. Current Methods of Measuring Exosystem Performance

A literature search was performed in Compendex, Inspec, NTIS, and PubMed with the terms "exoskeleton," "exosuit," "human," and "experiment," restricted to the years 2004 to 2017. Exoskeletons prior to this time period were described by Dollar and Herr [1]. This resulted in 1,240 unique papers. From these papers, we evaluated if a paper (1) was in English, (2) contained a human study ($n \geq 1$) in which exosystem hardware was worn, (3) if the exosystem spanned a lower extremity joint, and (4) if the exosystem had potential to be portable. This generated 146 papers describing 111 instances of exosystems. Of the papers removed, 42 were not in English and 5 were not accessible through interlibrary networks. The remaining removed papers included instances of literature reviews, prosthetic applications, signal processing methodologies without a human-system study, as well as upper extremity device development and evaluation. While we have highlighted systems designed to assist motor performance, exoskeletons in the literature review were included that augment, measure, and extract energy from motion. Table IV summarizes the joints augmented and the evaluation methods for the reviewed exosystems. We considered multiple references to a single exosystem as one instantiation of the system. For each exosystem, the degrees of freedom were categorized as passive or active assist. The study environments were dominated by overground and treadmill, but also included sit-to-stand, stairs, and in-air motions. Within "other" were balancing ($n = 1$), squatting ($n = 2$), load-bearing ($n = 4$), turning ($n = 3$), and an irregular surface ($n = 1$). Many human studies measure kinematics of the exosystem itself. Here the "other" measurements included hydraulic flow rate of the exoskeleton actuator ($n = 1$), power consumption ($n = 3$), muscle stiffness ($n = 1$), heart rate ($n = 1$), and EEG ($n = 1$).

TABLE IV
LOWER EXTREMITY EXOSYSTEM EVALUATION MEASUREMENTS. DESCRIPTION OF AND EVALUATION METHODS FOR 87 UNIQUE EXOSYSTEMS. COUNTS FOR A SINGLE EXOSYSTEM MAY APPEAR IN MORE THAN ONE TABLE CELL DUE TO TESTING OF MULTI-JOINT EXOSYSTEMS.

	Category	Hip	Knee	Ankle
Active/Passive	Active	52	82	29
	Passive	24	12	51
Study Environment	Overground	47	52	51
	Treadmill	20	27	21
	Stairs	6	8	7
	Sit-to-Stand	9	12	10
	In-Air Swing	8	15	9
	Other	11	12	12
Measurement Collected	Kinematics (motion capture)	11	15	14
	Kinematics (other)	47	72	56
	External Forces	34	39	34
	Joint Torques	24	33	24
	Surface Electromyography	18	31	25
	Metabolics	7	10	12
	Human-Exo Interaction	6	5	5
	Other	6	8	8

While many devices reported a type of kinematics data, there were fewer published reports for other measures. Within the category of measured kinematics, 22 exosystems used external motion capture to quantify the kinematics, while 80 exosystems determined kinematics using sensing onboard the exosystem or with another attached system. Only two studies disambiguated between motion of the human and motion of the exosystem. Among the devices tested overground, the modality was primarily straight and level locomotion. While there are studies in the biomechanics literature that highlight human performance in alternate environments (e.g., uneven terrain [70], stair negotiation [71], [72], avoiding obstacles [73], [74], and with perturbations [75], [76]), there is limited research on these environments within published exosystem evaluation.

From these results, we see that there is a need to expand the types of study environments and tasks to represent operational needs of the user. The fluency of the human-machine interaction becomes even more important as the human walking pattern becomes less regular, as the potential for more mode changes increases, and acceptable system reaction times may decrease.

While many papers did present kinematics, these data require expertise to understand and interpret. Depending on the operational goals, the interpretation of the kinematics may differ. For example, the joint angles during inclines or stair negotiation are expected to experience a different range of motion than level ground walking. Further, it may not be desired to replicate the nominal task biomechanics for all anthropometries, or in scenarios where the wearer is in a degraded state. There is still a need to develop operational performance metrics that will aid in evaluation of the varying system architectures. Further, an understanding of the difference in kinematics between the human and the exosystem may provide important information on the effects of misalignment or shifts in the system on performance. These needs highlight the importance of developing standard test methods.

B. Extending Measures of Exosystem Performance

Dollar and Herr [1] propose that systems that do not reduce the metabolic cost of the operator have very little value, and suggest that systems should reduce forces borne by the musculoskeletal system and improve bipedal stability. This description points to one potential class of exosystems. In the context of human systems integration, we propose that there are additional factors that must be considered, especially as other use cases for exosystems are developed. While reduction in metabolic energy may reduce fatigue and increase user endurance, this oversimplifies the exosystem operational problem statement. For example, a system may decrease metabolic consumption by reducing muscle activity in a specific joint (e.g., the ankle) while unintentionally increasing internal joint loads on a different joint (e.g., the hip) and thus potentially increase the risk of injury. Hence, although there may be an operational performance benefit in the short term, the long term consequences are not captured. On the other hand, if the total task time is decreased because of assistance from an exosystem, the overall power consumption of the human operator may be lower than manually completing a given task, despite a temporary increase in metabolic cost associated with use of the exosystem. At the time of Ferris et al. [12], they noted that only one study had reported oxygen consumption for powered walking. As seen in Table IV, more studies of lower extremity exosystems are including the measurement of metabolics. Still, it remains to be determined what additional metrics are appropriate based on the operational environment in which the exosystem is intended to be used.

While the measurement of body kinematics holds a vast wealth of understanding on the users' mobility, it is ultimately constrained by the tasks selected to perform. Earhart [77] considers locomotion of a person (without an exosystem) and highlights that current laboratory settings focus on short-term measures of gait during forward walking, which is not sufficient to understand dynamic postural control in the face of different environments, goals, biomechanical constraints, and sensory conditions. His suggestion of secondary tasks with non-locomotor demands is in-line with human factors considerations. A similar trend is found in the meta-analysis of exosystems performed here, in that current evaluation focuses on short-term measures of gait during forward walking either overground or on a treadmill. The evaluation of dynamic tasks and transitions between tasks, coupled with cognitive tasks, will be important for obtaining a systematic evaluation of exosystem performance that can transfer to operating environments. However, this additional task must be part of a comprehensive approach that includes computational modeling and simulation, laboratory testing, and field testing. Each component provides valuable information and builds on knowledge acquired to ensure that the system is ultimately successfully transitioned based on supporting performance data.

As highlighted in Section II-B, the use of an exosystem could require mental workload by the user, requiring the allocation of attentional resources, and has the potential for performance degradation in any additional tasks. There is also the possibility of irritation, chaffing, or thermal dissipation from the exosystem, increasing attention to the exosystem, thereby affecting system use and leading to performance degradation. There are several methods in the literature for measuring workload, including measures of primary task, secondary task(s), physiology, and subjective observer-based or self-reported evaluations [29], [78]. While primary task measures are simple to evaluate as they are already being collected, many times users can complete the primary task with the attentional resources they can allocate. Thus, a common method for analyzing workload is to include secondary tasks in different perceptual and processing modalities to find the thresholds at which performance degrades. Another option is to use a subjective test after use of the system, such as the NASA Task Load Index (TLX) [79], which allows users to rate their perceived task demands across workload dimensions. Ongoing efforts in using EEG to measure cognitive load [80] are interesting and may provide future methods for assessing workload with exosystems, although these methods are currently confounded by motion artifacts, are a superposition of multiple processes that need to be disambiguated, and are currently difficult to validate [81].

Measurements of SA using the operator's perception of their own awareness has been assessed using subjective scales [82]. However, such ratings are limited by the operator's own perceptions of reality and may correspond with confidence rather than SA. Similarly, ratings by expert observers [83] are constrained as observers may have only limited knowledge of the operators concept of the situation [84]. Objective measures are traditionally obtained via analysis of answers to questions administered during simulation freezes [85] or by measuring the reaction time to probe questions related to the displayed information [86]. However, these do not provide the ability to assess short-term temporal changes in critical SA variables and the experimenter must be careful about timing of questions to balance limitations in short term memory with task workload. To address these limitations, studies have analyzed the timeliness and accuracy of required verbal callouts [65] similar to those seen in aviation (e.g., [87], [88]). These methods of assessing task critical elements of SA have potential for exosystem evaluation. Each of these methods provides a way to assess SA, but must be interpreted in the context of their limitations. There is no way to currently map the cognitive perception, comprehension, and projection of a person objectively without explicit interaction.

In preparing to design and evaluate an exosystem, it is important to consider questions related to the operational use, physical factors, and cognitive factors. We highlight how the questions in operational use and cognitive factors can be considered for designers and evaluators. These lists are a suggested starting point to engage the reader in considering a human factors view of exosystem design and evaluation and is not meant to be an exhaustive list of all relevant questions.

Operational Use Case

1) In what environment will the system operate?
2) For what task(s) is the exosystem providing assistance?
3) What information or experience is required for synergistic (or satisfying) performance of the exosystem?
4) How long does it take to become an adapted user of the exosystem?

While the first three questions seem intuitive, these are important to discuss for designers as they are relevant to preparing the exosystem for improved training methods, as well as for generating design requirements. For the latter, the operational environment will drive potential hazards that must be accounted for (e.g., thermal constraints, shielding from dust). When considering training, feedback is used to enable both the building a mental model of the system (Section II-A) and for developing calibrated trust (Section II-C) in the system. Thus, at the design stages, we must consider what the appropriate environments and tasks will be such that the relevant sensor data or mode selection can be stored and presented to the user during training. Moulières-Seban et al. [5] provide a methodology for activity analysis, basic design, and detailed design that can aid in the engaging human factors concerns throughout the design process. Question 3 specifically highlights the consideration of the information that must be presented. Most current exosystems do not provide explicit feedback to the user, and we should ask ourselves if there is information that should be available for training or the operational use cases we select. Question 4 should be answered at the evaluation stage, through measures of relevant performance. Training methods that provide enhanced user feedback may decrease the time required to become adapted, as well as the selection of the control method as discussed previously. Adaptation to an exosystem may depend on the perceptual, motor, and decision making capabilities of the user, as well as the selected algorithms for the exosystem and the need to manually or autonomously tune any parameters. Studies in telerobotic manipulation are revealing the variability of human performance, the importance of algorithm selection on performance, as well as the style of training (e.g., [89]–[91]).

Cognitive Factors

1) Does the system require significant attentional resources to interact with or command?
2) How does interaction with the system during various nominal and off-nominal tasks affect SA?
3) What workload does the user experience when performing the functional task with the system?
4) Can a user detect relevant changes in mode and/or appropriately trigger changes in mode?
5) Are the "after-effects" when the user transitions between modes or takes off the system appropriately accounted for or minimized if detrimental?
6) Does the user have a mental model of the system dynamics and functional modes that enable proficiency?
7) Does the user trust the system to be safe and to act in accordance with his or her intentions?

The constructs of a mental model, attention, workload, and SA are difficult to measure. However, as described earlier in this section, there are methods to infer properties of these constructs. To address attention and workload, scenarios need to be constructed that will push the operator to a limit where more attentional resources are required than can be provided by the user. This may include having secondary or tertiary tasks, with the priorities of these tasks explicitly defined for the operator. The reduction in performance on the additional tasks is used to assess the differences in attentional resources. For example, if the exosystem is used by a warfighter, the relevant primary task may be a ruck march with a secondary task of radio communication. The surrogate for workload could be the time to respond to radio communications. For a civilian setting, the primary task may still be gait, with communication via cell phone as a secondary task. The objective measures of the secondary task can be paired with a subjective measure, such as the NASA TLX [79], to gain additional insight from the operator. To understand the user mental model and situation awareness, we recommend that specific questions be designed that are asked to the user. These questions should be defined for each level of SA to understand where any breakdowns occur (for examples, see Table V). With an understanding of where and for what aspect of the situation (e.g., exosystem or environment) breakdowns may occur in the evaluation, future system design requirements can be developed. Trust can also be evaluated with a set of questions through self-reported or observer-based measures [92]–[95] These heuristics were developed for unmanned systems and human-computer interactions scenarios and should be adapted by the exosystem community with an agreed-upon terminology such that the questions and actionable items are aligned with relevant exosystem taxonomies and use cases.

TABLE V
SAMPLE SITUATION AWARENESS QUESTIONS FOR EXOSYSTEM EVALUATION. HERE WE ASSUME THE TASK OF A WARFIGHTER ON PATROL.

SA Level	Sample Question
Related to Mode (Exosystem SA)	
1	What is the current operational mode of the exosystem?
2	What can the exosystem do in this mode? How would you shift to a different mode?
3	Given what you are currently doing, what will the exosystem do next?
Related to Scenario/Use Case (Environment SA)	
1	Did you see a person during your patrol?
2	Was the person you saw a friendly or enemy? Were they coming towards you or away from you?
3	Would this person intercept you on your current path?

Question 5 is important for system evaluators to consider in the context of the the use case. In a hospital rehabilitation setting, the presence of “after-effects” may be desired and encouraged, whereas they could be detrimental during soldier engagement. Transitions between modes needs to be considered in the context of these after-effects and the enabling of system trust, which enables proficiency.

Addressing these questions will require a deeper understanding of the operational requirements of the system. If we consider transportation, sometimes we need a car and sometimes we need a long-haul truck with multiple trailers. Each of these vehicles is selected based on features (desired qualities) relevant to the needs imposed by the required task. In designing exosystems, it must be made explicit what the requirements are for a particular system so that it may be evaluated with the appropriate verification and validation standards. A heavy-lift enhancing exosystem for use in a depot may have requirements on balance and strength, whereas an enhancing exosystem for use during soldier engagement may have requirements on speed and agility. There are underlying tradeoffs in these higher level metrics; for example, a system that has high agility may not have as high a balance score. Stirling and MacLean [96] describe a methodology for characterizing the subject matter, defining relevant information, and formalizing requirements for generating and assessing performance metrics for Occupational Therapy decision making aids. Similar to this evaluation, the exosystem community (e.g., designers, evaluators, and end-users) must articulate the higher-level feature needs such that performance metrics can be generated that align with these relevant features and permit a set of verification and validation requirements to be defined. By better understanding and quantifying the interactions between system design parameters and human factors principles and practices, we believe exosystems can more readily be adopted and used in operational environments.

IV. Conclusions and Future Work

There have been decades of work in the domain of human factors and human-system integration in applications for complex systems to understand concepts within physical domains (e.g. system sizing and injury risk) and cognitive domain (e.g., workload, situation awareness, and automation mode confusion). Exosystem technologies are reaching the point where the transition to an operational environment is near-term feasible. Rather than recreate our understanding of human-system integration, we should learn from the literature and use our current understanding to better design and evaluate exosystems for use in operational environments. In this paper, cognitive human factors were applied to exosystem applications and important questions to consider when designing and evaluating exosystems were posed.

Acknowledgment

The authors would like to thank Ms. Tingxiao Sun for her help in categorizing the literature review papers and Prof. Nancy Leveson for her comments during the preparation of this manuscript. This work was supported in part by NSF Award IIS-1453141 (LS), the Jeptha and Emily V. Wade Award (HS, LS), and Draper internal opportunity investment funding (KD, EJ). The views expressed in the manuscript are not endorsed by the sponsors.

References

[1] A. M. Dollar and H. Herr, “Lower extremity exoskeletons and active orthoses: Challenges and state-of-the-art,” *IEEE Transactions on Robotics*, vol. 24, no. 1, pp. 144–158, 2008.

[2] B. Chen, H. Ma, L.-Y. Qin, F. Gao, K.-M. Chan, S.-w. Law, L. Qin, and W.-H. Liao, “Recent developments and challenges of lower extremity exoskeletons,” *Journal of Orthopaedic Translation*, vol. 5, pp. 26–37, 2016.

[3] K. Anam and A. A. Al-Jumaily, "Active exoskeleton control systems: State of the art," *Procedia Engineering*, vol. 41, pp. 988–994, 2012.

[4] T. Jacobs and V. G. Singh, "ISO 13482 The new safety standard for personal care robots," in *International Symposium on Robotics*, 2014, pp. 698–703.

[5] T. Moulières-Seban, D. Bitonneau, J.-M. Salotti, J.-F. Thibault, and B. Claverie, *Human Factors Issues for the Design of a Cobotic System*. Cham: Springer International Publishing, 2017, pp. 375–385.

[6] D. K. Mitchell, "Mental workload and arl workload modeling tools," DTIC Document, Tech. Rep., 2000.

[7] S. Yee, L. Nguyen, P. Green, J. Oberholtzer, and B. Miller, "Visual, auditory, cognitive, and psychomotor demands of real in-vehicle tasks," University of Michigan, Ann Arbor, Transportation Research Institute, Tech. Rep. UMTRI-2006-20, 2007.

[8] K. E. Gordon, C. R. Kinnaird, and D. P. Ferris, "Locomotor adaptation to a soleus EMG-controlled antagonistic exoskeleton." *Journal of neurophysiology*, vol. 109, no. 7, pp. 1804–14, apr 2013.

[9] S. Galle, P. Malcolm, W. Derave, and D. De Clercq, "Adaptation to walking with an exoskeleton that assists ankle extension," *Gait and Posture*, vol. 38, no. 3, pp. 495–499, 2013.

[10] K. R. Boff, "Revolutions and shifting paradigms in human factors & ergonomics," *Applied Ergonomics*, vol. 37, pp. 391–399, 2006.

[11] D. B. Chaffin, G. Andersson, and B. J. Martin, *Occupational Biomechanics, Fourth Edition*. Wiley, 2006.

[12] D. P. Ferris, G. S. Sawicki, and M. a. Daley, "A Physiologist'S Perspective on Robotic Exoskeletons for Human Locomotion," *International Journal of Humanoid Robotics*, vol. 04, no. 3, pp. 507–528, 2007.

[13] T. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*. The MIT Press, 1992.

[14] R. R. Burridge, A. A. Rizzi, and D. E. Koditschek, "Sequential composition of dynamically dexterous robot behaviors," *The International Journal of Robotics Research*, vol. 18, no. 6, pp. 534–555, 1999.

[15] N. Leveson, L. D. Pinnel, S. D. Sandys, S. Koga, and J. D. Reese, "Analyzing software specifications for mode confusion potential," in *Proceedings of a Workshop on Human Error and System Development*, 1997, pp. 132–146.

[16] H. A. Varol, F. Sup, and M. Goldfarb, "Multiclass real-time intent recognition of a powered lower limb prosthesis," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 3, pp. 542–551, 2010.

[17] N. Sarter, "Investigating Mode Errors on Automated Flight Decks : Illustrating the Problem-Driven , Cumulative , and Interdisciplinary Nature of Human Factors Research," *Human Factors*, vol. 50, no. 3, pp. 506–510, 2008.

[18] S. Wood, "Flight crew reliance on automation," Civil Aviation Authority, Tech. Rep. CAA report no. 2004/10, 2004.

[19] D. C. May, K. J. Holler, C. L. Bethel, L. Strawderman, D. W. Carruth, and J. M. Usher, "Survey of Factors for the Prediction of Human Comfort with a Non-anthropomorphic Robot in Public Spaces," *International Journal of Social Robotics*, vol. 9, no. 2, pp. 165–180, 2017.

[20] N. Jarrasse, M. Maestrutti, G. Morel, and A. Roby-Brami, "Robotic prosthetics," *IEEE Technology and Society Magazine*, vol. 34, no. 2, pp. 71–79, 2015.

[21] J. R. Lackner and P. Dizio, "Rapid adaptation to coriolis force perturbations of arm trajectory," *Journal of Neurophysiology*, vol. 72, no. 1, pp. 299–313, 1994.

[22] R. Shadmehr and F. a. Mussa-Ivaldi, "Adaptive representation of dynamics during learning of a motor task." *The Journal of neuroscience*, vol. 14, no. 5, pp. 3208–3224, 1994.

[23] R. Scheidt, D. Reinkensmeyer, M. Conditt, W. Rymer, and F. Mussa-Ivaldi, "Persistence of motor adaptation during constrained, multi-joint, arm movements." *Journal of neurophysiology*, vol. 84, no. 2, pp. 853–862, 2000.

[24] I. Cajigas, M. T. Goldsmith, A. Duschau-Wicke, R. Riener, M. A. Smith, E. N. Brown, and P. Bonato, "Assessment of lower extremity motor adaptation via an extension of the Force Field Adaptation Paradigm," *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, pp. 4522–4525, 2010.

[25] K. E. Gordon and D. P. Ferris, "Learning to walk with a robotic ankle exoskeleton," *Journal of Biomechanics*, vol. 40, no. 12, pp. 2636–2644, 2007.

[26] G. Baroni, A. Pedrocchi, G. Ferrigno, J. Massion, and A. Pedotti, "Static and dynamic postural control in long-term microgravity: evidence of a dual adaptation," *Journal of Applied Physiology*, vol. 90, no. 1, pp. 205–215, 2001.

[27] A. Dix, J. Finlay, G. Abowd, and R. Beale, *Human-Computer INteraction*, 3rd ed. Pearson Prentice Hall, 2004.

[28] J. A. Jacko, Ed., *The Human Computer Interaction Handbook: Fundamentals, Evolving Technolgies, and Emerging Applications*, 3rd ed. CRC Press, 2012.

[29] R. Proctor and T. Van Zandt, *Human Factors in Simple and Complex Systems, Second Edition*. CRC Press Taylor & Francis Group, 2008.

[30] G. Salvendy, Ed., *Handbook of human factors and ergonomics*. John Wiley & Sons, 2012.

[31] C. D. Wickens, J. G. Hollands, S. Banbury, and R. Parasuraman, *Engineering psychology & human performance*. Psychology Press, 2015.

[32] W. B. Rouse and N. M. Morris, "On looking into the black box: Prospects and limits in the search for mental models," Office of Naval Research, Arlington, VA, Tech. Rep. AD-A159 080, 1985.

[33] J. Rasmussen, "Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, no. 3, pp. 257–266, 1983.

[34] D. Norman, "Some observations on mental models," in *Mental Models*, D. Gentner and A. Stevens, Eds. Psychology Press, 1983, ch. 1, pp. 7–14.

[35] W. Cornwall, "In pursuit of the perfect power suit," *Science*, vol. 350, no. 6258, pp. 270–273, 2015.

[36] J. Y. Chen, M. J. Barnes, and M. Harper-Sciarini, "Supervisory control of multiple robots: Human-performance issues and user-interface design," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 4, pp. 435–454, 2011.

[37] C. Hochberg, O. Schwarz, and U. Sceneider, "Aspects of Human Engineering – Bio-optimized Design of Wearable Machines," in *Soft Robotics: Transferring Theory to Application*, A. Verl, A. Albu-Schaffer, O. Brock, and A. Raazt, Eds. Springer, 2015, ch. 16, pp. 184–197.

[38] J. A. Macedo, D. B. Kaber, M. Endsley, P. Powanusorn, and S. Myung, "The Effect of Automated Compensation for Incongruent Axes on Teleoperator Performance," *Human Factors*, vol. 40, no. 4, pp. 541–553, 1998.

[39] J. Zhang, C. C. Cheah, and S. H. Collins, "Experimental comparison of torque control methods on an ankle exoskeleton during human walking," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 5584–5589.

[40] Y. Ding, F. A. Panizzolo, C. Siviy, P. Malcolm, I. Galiana, K. G. Holt, and C. J. Walsh, "Effect of timing of hip extension assistance during loaded walking with a soft exosuit," *Journal of neuroengineering and rehabilitation*, vol. 13, no. 1, p. 87, 2016.

[41] D. Kahneman, *Attention and Effort*. Prentice Hall, Inc., 1973.

[42] C. T. Connolly and G. Tenenbaum, "Exertion-Attention-Flow linkage under different workloads," *Journal of Applied Social Psychology*, vol. 40, no. 5, pp. 1123–1145, 2010.

[43] M. Endlsey, "Toward a theory of situation awareness in dynamic systems," *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995.

[44] C. D. Wickens, "Situation Awareness and Workload in Aviation," *Current directions in psychological science*, vol. 11, no. 4, pp. 128–133, 2002.

[45] A. D. Dragan, K. C. T. Lee, and S. Srinivasa, "Legibility and Predictability of Robot Motion," in *ACM/IEEE International Conference on Human Roboti Interaction*, Tokyo, Japan, 2013, pp. 301—-308.

[46] S. Tellex, R. Knepper, A. Li, D. Rus, and N. Roy, "Asking for help using inverse semantics," in *Proceedings of Robotics: Science and Systems*, Berkeley, USA, July 2014.

[47] A. Powers and S. Kiesler, "The advisor robot: Tracing people's mental model from a robot's physical attributes," in *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction*, ser. HRI '06. New York, NY, USA: ACM, 2006, pp. 218–225. [Online]. Available: http://doi.acm.org/10.1145/1121241.1121280

[48] K. Baraka, S. Rosenthal, and M. Veloso, "Enhancing Human Understanding of a Mobile Robot ' s State and Actions using Expressive Lights," in *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. Columbia University, NY, USA: IEEE, 2016.

[49] L. Marchal-Crespo and D. J. Reinkensmeyer, "Review of control strategies for robotic movement training after neurologic injury," *Journal of Neuroengineering and Rehabilitation*, vol. 6, 2009.

[50] R. Parasuraman, T. Sheridan, and C. Wickens, "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man, and Cybernetics — Part A: Systems and Humans*, vol. 3, no. 30, pp. 286–297, 2000.

[51] J. Hart and J. Valasek, "Methodology for prototyping increased levels of automation for spacecraft rendezvous functions," in *AIAA Conference and Exhibit*, 2007.

[52] M. Gombolay, A. Bair, C. Huang, and J. Shah, "Computational design of mixed-initiative human robot teaming that considers human factors : situational awareness , workload , and workflow preferences," *International Journal of Robotic Research*, vol. 36, no. 5-7, pp. 597–617, 2017.

[53] M. A. Goodrich and A. C. Schultz, "Human Robot Interaction : A Survey," *Foundations and Trends in Human-Computer Interaction*, vol. 1, no. 3, pp. 203–275, 2007.

[54] P. Scerri, D. Pynadath, and M. Tambe, "Towards Adjustable Autonomy for the Real World," *Journal of Artificial Intelligence Research*, vol. 17, pp. 171–228, 2002.

[55] M. C. Gombolay, R. A. Gutierrez, S. G. Clarke, G. F. Sturla, and J. A. Shah, "Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams," *Autonomous Robots*, vol. 39, no. 3, pp. 293–312, Oct 2015. [Online]. Available: https://doi.org/10.1007/s10514-015-9457-9

[56] J. Baraglia, M. Cakmak, Y. Nagai, R. Rao, and M. Asada, "Initiative in Robot Assistance during Collaborative Task Execution," in *ACM/IEEE International Conference on Human Robot Interaction*, Christchurch, New Zealand, 2016, pp. 67–74.

[57] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human Performance Issues and User Interface Design for Teleoperated Robots Human Performance Issues and User Interface Design for Teleoperated Robots," *IEEE Transactions on Systems, Man, and Cypernetics - Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1231–1245, 2007.

[58] M. Wehner, B. Quinlivan, P. M. Aubin, E. Martinez-Villalpando, M. Baumann, L. Stirling, K. Holt, R. Wood, and C. Walsh, "A lightweight soft exosuit for gait assistance," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2013, pp. 3362–3369.

[59] P. Malcolm, W. Derave, S. Galle, and D. De Clercq, "A Simple Exoskeleton That Assists Plantarflexion Can Reduce the Metabolic Cost of Human Walking," *PLoS ONE*, vol. 8, no. 2, pp. 1–7, 2013.

[60] S. Galle, P. Malcolm, W. Derave, and D. De Clercq, "Enhancing performance during inclined loaded walking with a powered ankle-foot exoskeleton." *European journal of applied physiology*, vol. 114, no. 11, pp. 2341–51, 2014.

[61] C.-m. Huang and B. Mutlu, "Anticipatory Robot Control for Efficient Human-Robot Collaboration," in *ACM/IEEE International Conference on Human-Robot INteraction*, Chrischurch, New Zealand, 2016.

[62] J. Mainprice and D. Berenson, "Human-Robot Collaborative Manipulation Planning Using Early Prediction of Human Motion," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013.

[63] K. P. Hawkins and A. F. Bobick, "Anticipating human actions for collaboration in the presence of task and sensor uncertainty," in *IEEE International Conference on Robotics and Automation*, Hong Kong, China, 2014, pp. 2215–2222.

[64] N. B. Sarter, D. D. Woods, and C. E. Billings, "Automation surprises," 1997, vol. 2, pp. 1926–1943.

[65] C. J. Hainley, K. R. Duda, C. M. Oman, and A. Natapoff, "Pilot Performance, Workload, and Situation Awareness During Lunar Landing Mode Transitions," *Journal of Spacecraft and Rockets*, vol. 50, no. 4, pp. 793–801, 2013.

[66] J. Lee and K. See, "Trust in automation: Designing for appropriate reliance," *Human Factors*, vol. 46, no. 1, pp. 50–80, 2004.

[67] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors*, vol. 39, no. 2, pp. 230–253, 1997.

[68] P. Madhavan and D. Wiegmann, "Similarities and differences between human-human and human-automation trust: an integrative review," *Theoretical Issues in Ergonomics Science*, vol. 8, no. 4, pp. 277–301, 2007.

[69] N. Leveson, *Engineering a Safer World: Applying Systems Thinking to Safety*. MIT Press, 2012.

[70] A. S. Voloshina, A. D. Kuo, M. a. Daley, and D. P. Ferris, "Biomechanics and energetics of walking on uneven terrain." *The Journal of Experimental Biology*, vol. 216, pp. 3963–70, 2013.

[71] L. A. Livingston, J. M. Stevenson, and S. J. Olney, "Stairclimbing kinematics on stairs of differing dimensions," *Arch Phys Med Rehabil*, vol. 72, no. 6, pp. 398–402, 1991.

[72] R. Riener, M. Rabuffetti, and C. Frigo, "Stair ascent and descent at different inclinations," *Gait & posture*, vol. 15, no. 1, pp. 32–44, 2002.

[73] J. Krell and A. Patla, "The influence of multiple obstacles in the travel path on avoidance strategy," *Gait & posture*, vol. 16, no. 1, pp. 15–19, 2002.

[74] A. Patla and M. Greig, "Any way you look at it, successful obstacle negotiation needs visually guided on-line foot placement regulation during the approach phase," *Neuroscience Letters*, vol. 397, no. 1-2, pp. 110–114, 2006.

[75] L. M. Nashner, "Balance adjustments of humans perturbed while walking," *Journal of Neurophysiology*, vol. 44, no. 4, pp. 650–664, 1980.

[76] M. H. Woollacott and P.-F. Tang, "Balance control during walking in the older adult: research and its implications," *Physical therapy*, vol. 77, no. 6, pp. 646–660, 1997.

[77] G. Earhart, "Dynamic Control of Posture Across Locomotor Tasks," *Movement Disorders*, vol. 28, no. 11, pp. 1501–1508, 2013.

[78] C. Wickens, J. Lee, Y. Liu, and S. Gordon Becker, *An Introduction to Human Factors Engineering, Second Edition*. Pearson Prentice Hall, 2004.

[79] S. Hart and L. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," in *Human Mental Workload*, P. Hancock and N. Meshkati, Eds. North Holland Press, 1988.

[80] C. Berka, D. J. Levendowski, M. N. Lumicao, A. Yau, G. Davis, V. T. Zivkovic, R. E. Olmstead, P. D. Tremoulet, and P. L. Craven, "Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviation, space, and environmental medicine*, vol. 78, no. Supplement 1, pp. B231–B244, 2007.

[81] P. Hancock and J. Szalma, "The future of neuroergonomics," *Theoretical Issues in Ergonomics Science*, vol. 4, no. 1-2, pp. 238–249, 2003.

[82] R. Taylor, "Situation awareness rating technique (SART): The development of a tool for aircrew systems design," in *Proceedings of the AGARD Symposium on Situational Awareness in Aerospace Operations: Aerospace Medical Panel Symposium*, 1990.

[83] Y. Hauss and K. Eyferth, "Securing future atm-concepts safety by measuring situation awareness in atc," *Aerospace Science and Technology*, vol. 7, no. 6, pp. 417–427, 2003.

[84] M. Endlsey and D. Kaber, "Level of automation effects on performance, situation awareness and workload in a dynamic control task," *Ergonomics*, vol. 42, no. 3, pp. 462–492, 1999.

[85] M. Endsley, "Situation awareness global assessment technique (SAGAT)," in *Proceedings of the National Aerospace and Electronics Conference*, 1988, pp. 789–795.

[86] F. T. Durso and S. D. Gronlund, "Situation awareness," in *Handbook of Applied Cognition*, F. T. Durso, R. S. Nickerson, and R. W. Schvaneveldt, Eds. John Wiley and Sons Ltd., 1999, ch. 10, pp. 283–314.

[87] R. Machado, "Cockpit conversations: Self talk for survival," in *Rod Machado's Plane Talk: The Mental Art of Flying an Airplane, Third Edition*, R. Machado, Ed. The Aviation Speaker's Bureau, 2003, ch. 3, pp. 3/1–14.

[88] P. Bergqvist, "Talk to yourself," *Flying*, 2011.

[89] M. Talha, A. M. E. Ghalamzan, C. Takahashi, J. Kuo, W. Ingamells, and R. Stolkin, "Towards robotic decommissioning of legacy nuclear plant : results of human-factors experiments with tele-robotic manipulation , and a discussion of challenges and approaches for decommissioning," in *IEEE International Symposium on Safety, Security, and Rescue Robots*, Lausanne, Switzerland, 2016, pp. 166–173.

[90] R. Forman, "Objective Performance Metrics for Improved Space Telerobotics Training," Master of Science, Massachusetts Institute of Technology, 2011.

[91] N. Paperno, M. Rupp, E. M. Maboudou-tchao, J. A.-a. Smither, and A. Behal, "A Predictive Model for Use of an Assistive Robotic Manipulator : Human Factors Versus Performance in Pick-and-Place / Retrieval Tasks," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 6, pp. 846–858, 2016.

[92] K. Jackson, Z. Prasov, E. Vincent, and E. Jones, "A heuristic based framework for improving design of unmanned systems by quantifying and assessing operator trust," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 60, no. 1, 2016, pp. 1696–1700.

[93] K. E. Schaefer, D. R. Billings, J. L. Szalma, J. K. Adams, T. L. Sanders, J. Y. Chen, and P. A. Hancock, "A meta-analysis of factors influencing the development of trust in automation: Implications for human-robot interaction," Army Research Laboratory, Tech. Rep. ARL-TR-6984, 2014.

[94] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, "Foundations for an empirically determined scale of trust in automated systems," *International Journal of Cognitive Ergonomics*, vol. 4, no. 1, pp. 53–71, 2000.

[95] J. Meyer and J. D. Lee, *Trust, Reliance, and Compliance*. Oxford University Press, 2013.

[96] L. Stirling and J. Maclean, "Roadmap for the development of at-home telemonitoring systems to augment occupational therapy," *IEEE Transactions on Human-Machine Systems*, vol. 4, no. 46, pp. 569 – 580, 2016.