Introduction: People who are influenced by changing health or environmental factors are not consistently reliable when operating complex machines. While existing research seeks to measure a human's trust in their robotic partner[1-3] or discover novel means of human-robot control-sharing[4], little quantifies the extent to which a robot should act on its human's commands. Without a measure of robotic trust, the robot cannot temper its actions and prevent undesired or dangerous outcomes due to lacking human cogency or ability. Rehabilitation and assistive robotics are particularly salient domains for studying robotic trust because human input must be reliable for continued mobility. The aim of this proposal is twofold: first, *to define an effective measure of robotic trust* and second, *to evaluate different methods of applying robotic trust within the domain of human-robot control sharing.*

Aim 1: Identify effective measures of robotic trust based on human inputs. Three initial definitions of robotic trust will be analyzed: **divergence, comprehension,** and **performance**. Inspired by [5], which defined robotic trust as the difference between optimal control trajectories and a human control inputs, **divergence** will inspect the difference between the human's control inputs and the robot's trajectories. Computations of divergence may be simple (e.g. the dot product) or complex (e.g. the Fréchet distance[5]). **Comprehension** will quantify the human's understanding of their robotic partner's physical limitations via control theoretic system stability metrics[6-7] and safety policy violations (e.g. frequency of autonomous interventions[8]). Finally, **performance** uses task-agnostic means (e.g. efficiency measures such as mean completion time[9]) to quantify the human's ability to effectively accomplish an assigned task. An effective trust measurement may be a combination of all three of the aforementioned measures. Combinations of the methods may aggressively select the best performing trust metric, temporally average trust metrics, or weigh trust metrics in a blending scheme where weights are learned to maximize overall trust for an individual or task.

Aim 2: Applications of robotic trust measures to improve shared control. Recent trends in assistive and rehabilitation robotics seek to maximize the human's control of their robotic partner while simultaneously ensuring safety and efficiency[10-11]. In **linear autonomy blending**[8], trust may directly allocate the autonomy blending parameter combining human and robot control. In **autonomy allocation**, the thresholds on measured human control signals may be dynamically adjusted by robotic trust to define when to shift between discrete autonomy levels[12] (e.g. obstacle avoidance, full autonomy, etc.). Recent literature indicates the need to break some tasks into subtask primitives[4,13-14] for control sharing. This suggests development of **task-dependent trust** measures. Intent inference models[15] can be used to predict the human's task, break it into subtasks, and dynamically shift between appropriate measures of robotic trust.

Evaluation: Trust measures will be tested on commercial robotic arms and wheelchair platforms with both healthy and differently-abled populations. Comparing explicitly-calculated robotic trust measures to task-specific performance measures will help determine which robotic trust measures are most effective (aim 1). Comparing user performance and preferences, with and without robotic assistance, will inform which applications of robotic trust are most effective (aim 2). Prior to robotic platform evaluations, subjects will validate robotic trust measures and applications through robotic simulations built using Gazebo[16]. The widely used NASA Multi-Attribute Task Battery[17] "distraction task" will be used to vary task difficulty via cognitive load modulation in both simulation and hardware testing.

Specific Resources: The argallab at Northwestern University has a custom, autonomous wheelchair platform and two robotic arm platforms from Kinova Robotics[18], each of which can be used to validate various trust formulations on real hardware. Furthermore, the argallab is situated within the Shirley Ryan AbilityLab, the United States' premier rehabilitation hospital. This facility permits robotic trust studies with numerous patient populations and collaborations with world-leading clinicians. Moreover, a large majority of recent research in robotic trust has been conducted at Northwestern University by argallab Director Prof. Brenna Argall and collaborator Prof. Todd Murphey, allowing for access to expert mentorship and insight.

Intellectual Merit: This research is a stepping stone towards developing robots that can differentiate useful human commands from those that are dangerous or noisy due to the individual's unique control capabilities. Few methods presently exist that quantify the human's cogency when issuing commands in the human-robot team. The proposed robotic trust research will investigate and define multiple, validated methods to quantify the quality of the human's commands as derived from the human's behavior. Tempering existing shared-control frameworks with robotic trust will develop a model capable of subtask discretization and define task-specific robotic trust measures.

Broader Impacts: Many people within differently-abled communities struggle to control the very devices designed to improve their quality of life and maintain their freedom of mobility. This research will more appropriately allocate assistance within the human-robot team and enable robots to adapt to their partner's ever-changing abilities. It is hypothesized that this adaptation will improve the safety and performance of the human-robot team in not only assistive and rehabilitation robotics, but also in manufacturing and the automotive industry. Finally, this research has significant military applications; when a human becomes distracted in a warfighter scenario, the online adaptation of their robotic partner may shift the outcome of a life-or-death scenario.

Communication, Dissemination, and Outreach: Annual demonstrations during National Robotics Week at the Museum of Science and Industry in Chicago, an event mostly attended by middle and high school students, will communicate this research to a wider audience. To maintain their curiosity and break the monotony of extended hospital stays, the pediatric inpatients (aged 6-18) at the Shirley Ryan AbilityLab will be presented with regular opportunities to participate in STEM demonstrations using the robotic platforms designed for this research. Finally, to inspire the next generation of roboticists, undergraduate students in the argallab will work with and be mentored by graduate students on robotic trust-related projects.

Works Cited: **1.** M. Chen *et al. Proc. of HRI*. 2018. 307-315; **2.** C. Basu *et al. Proc. of AAAI Spring. Symp.* 2016. 85-91; **3.** H. Saeidi *et al. Proc. of IEEE Conf. on CDC.* 2015. 6052-6057; **4.** S.Music *et al. Annual Reviews in Control*. 2017. 44: 342-354; **5.** A. Broad *et al*. *IEEE RAL*, 2017. 2-1: 239-246; **6.** G. Franklin *et al. Feedback Control*. Pearson. 2014; **7.** X. Xu *et al. Proc. of IFAC on ADHS.* 2015. 48-27: 54-61; **8.** A. Erdogan *et al. Robotics and Autonomous Systems*, 2017. 94: 282-297; **9.** A. Steinfeld *et al. Proc of HRI.* 2006; **10.** O. Horn. *Proc. of ICSCS.* 2012. 1-6; **11.** D. Gopinath *et al. IEEE RAL*, 2017. 2-1: 247-254; **12**. M. Chiou *et al. Proc. of IROS.* 2016. 3581-3588; **13**. X. Yang *et al. Proc. of ICRA*. 2017. 5948-5953; **14.** M. Young, **C. Miller**, *et al. Proc. of ICRA.* 2019. (In-Review); **15**. S. Javdani *et al. IJRR*. 2018. 37-7:717-742; **16.** "Gazebo." gazebosim.org. [Oct. 10, 2018]; **17.** S. Hart *et al. Advances in Psych.* 1988. 52: 139-183; **18.** "Kinova." www.kinovarobotics.com [Oct. 15, 2018].