**Human Response to Robotic Failure**

An increasing global trend is the inclusion of autonomous and robotic systems in environments that previously lacked such methods. The use of autonomous and robotic systems is proliferated by hobbyists, researchers, and businesses. Potential risks include interference from individuals that feel threatened, unintentional violation of laws such as occupying illegal air space, and system malfunction. System malfunction can be costly. Does a robot’s malfunction rate or failure rate affect how a human interacts with a robot? How does robotic failure or abnormal behavior affect how a human interacts with the robot? Comments on end user interaction with robotic systems that exhibit failure when present are often a small part of the description of robotic system testing. A knowledge of how humans interact with autonomous and robotic systems that exhibit failure will aid developers in designing robust systems.



Table 1: Papers with applicable machine (automata, unmanned ground vehicle, unmanned aerial vehicle) types, cited metrics for both humans and machines, and the corresponding conclusions of which factors impact human-machine trust.

High failure rates of autonomous and robotic systems have a negative effect on human robot interactions when the failures cause injury or death such as failures involving surgical robots [17]. An exception to this is autonomous and robotic systems employed by military personnel. Military personnel are perceived as having blind trust in computer systems even after failures that cause unintentional human casualties [16]. Competition objective failures of robots such as the extent of robot mobility forces operators to be less exploratory, increases operator stress, and increases operator fear of further failure resulting in the operator not attempting specific competition events [1]. Incorporating feedback of a machine’s confidence in a decision made or a machine’s component has a positive effect on operator trust [8], [11]. Table 1 shows the factors that affect appropriate human-machine trust while varying a subset of human and/or robot dependent variables.



Table 2. Metrics cited for machines and humans and their relative frequencies grouped by machine type.

The factors affecting human-robot trust are divisible into machine metrics and human metrics. This division is represented in table 2 with the most prominent metrics cited. Machine metrics include: mean time between failures (MTBF); availability of the machine for use; performance during an event; task, mission, or test; predictability of machine response; amount and predictability of false positives; levels of automation; adaptability to changing conditions; system safety; accuracy of reported information such as system state; and impact of machine failure on long term mission performance. Isolated occurrences of false positives has caused the destruction of civilian craft by military personnel and resulted in tragic loss of life [16]. Increased levels of automation provide additional software paths for failures to occur while being able to switch between automation levels seamlessly increases operator trust [3], [9], [13]. Safety considerations factor into machines used to assist disabled individuals as well as systems used in unsafe or potentially dangerous situations [8], [13]. Accurate confirmable reported data increases operator trust in the applicable system [3]. Human metrics include: preconceived expectations of robot capabilities, an individual’s propensity to trust, system expertise, an individual’s self-confidence in their own performance, the time required for a human to repair a failed component (Mean Time To Repair), operator workload, operator situational awareness, and an individual’s propensity to take risks. Preconceived expectations of a machine’s capabilities are often shaped by comparative fictional examples of the applicable machine [6] or by mission specifications [16]. System expertise coupled with MTBF will affect an operator’s baseline trust of a machine. Increased MTBFs and increased expertise results in increased operator trust while decreased MTBFs and increased expertise reduces operator trust in the system [3]. An operator’s self-confidence affects total errors of the human-machine system as well as an operator’s desired reliance on machine automation versus their own abilities [6], [9]. The MTTR a system affects the systems overall availability and an operators trust in the system being available for use [7].

There are plenty of ways to measure the trust between humans and machines. Though, human expertise and expectations as well as machine performance and MTBF are the most common metrics studied across differing machine types that include automata, unmanned ground vehicles, and unmanned aerial vehicles. The addition of accurate machine confidence levels will improve overall human-machine trust. Knowing there is a measure of error inherent in a calculation or component will allow an operator to apply appropriate trust to a machines actions.

REFERENCE

[1]B. A. Maxwell, W. Smart, A. Jacoff, J. Casper, B. Weiss, J. Scholtz, H. Yanco, M. Micire, A. Stroupe, D. Stormont, and T. Lauwers, “2003 AAAI Robot Competition and Exhibition,” AI Magazine, vol. 25, no. 2, p. 68, Jun. 2004.

[2]R. Murphy, D. Shell, A. Guerin, B. Duncan, B. Fine, K. Pratt, and T. Zourntos, “A Midsummer Night’s Dream (with flying robots),” Auton Robot, vol. 30, no. 2, pp. 143–156, Oct. 2010.

[3]J. M. Oglesby, K. Stowers, K. Leyva, A. Dietz, S. Sonesh, S. Burke, and E. Salas, “Assessing Human-Automation System Safety, Efficiency, and Performance Developing a Metrics Framework,” Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 58, no. 1, pp. 1149–1153, Sep. 2014.

[4]C. Flanagin, “A Survey of Robotics Systems and Performance Analysis - robots.pdf.” [Online]. Available: http://www.cse.wustl.edu/~jain/cse567-11/ftp/robots.pdf. [Accessed: 18-Nov-2015].

[5]M. Desai, K. Stubbs, A. Steinfeld, and H. Yanco, “Creating trustworthy robots: Lessons and inspirations from automated systems,” 2009.

[6]M. Desai, S. Gadea-Omelchenko, C. Bruggerman, A. Steinfeld, H. Yanco, M. Medvedev, M. Vazquez, and S. McSheehy, “Effects\_of\_Changing\_Reliability\_on\_Trust\_of\_Robot\_Systems.pdf.” [Online]. Available: http://robotics.cs.uml.edu/fileadmin/content/publications/2012/Effects\_of\_Changing\_Reliability\_on\_Trust\_of\_Robot\_Systems.pdf. [Accessed: 02-Dec-2015].

[7]J. Carlson, R. R. Murphy, and A. Nelson, “Follow-up analysis of mobile robot failures,” in 2004 IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA ’04, 2004, vol. 5, pp. 4987–4994 Vol.5.

[8]Q. A. Do Hoang, J. Guiochet, D. Powell, and M. Kaaniche, Human-robot interactions: model-based risk analysis and safety case construction. .

[9]J. Y. Chen, E. C. Haas, K. Pillalamarri, and C. N. Jacobson, “Human-Robot Interface: Issues in Operator Performance, Interface Design, and Technologies,” Jul. 2006.

[10]V. Groom, L. Takayama, P. Ochi, and C. Nass, “I Am My Robot: The Impact of Robot-building and Robot Form on Operators.” [Online]. Available: http://web.cs.wpi.edu/~gogo/hive/papers/GroomEtAl\_2009\_HRI.pdf. [Accessed: 17-Nov-2015].

[11]M. Desai, P. Kaniarasu, M. Medvedev, A. Steinfeld, and H. Yanco, “Impact of robot failures and feedback on real-time trust,” in 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2013, pp. 251–258.

[12]A. Xu and G. Dudek, “isrr2013\_trust\_study.pdf.” [Online]. Available: http://www.cim.mcgill.ca/%7emrl/pubs/anqixu/isrr2013\_trust\_study.pdf. [Accessed: 18-Nov-2015].

[13]E. J. Vanderperre and S. S. Makhanov, “Overall availability and risk analysis of a general robot–safety device system,” International Journal of Systems Science, vol. 46, no. 10, pp. 1889–1896, Jul. 2015.

[14]P. Kaniarasu, A. Steinfeld, M. Desai, and H. Yanco, “Potential measures for detecting trust changes,” in Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, 2012, pp. 241–242.

[15]P. Kaniarasu, A. Steinfeld, M. Desai, and H. Yanco, “Robot confidence and trust alignment,” in Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction, 2013, pp. 155–156.

[16]P. W. Singer, “Robots at War: The New Battlefield.” [Online]. Available: http://wilsonquarterly.com/stories/robots-at-war-the-new-battlefield/. [Accessed: 10-Nov-2015].

[17]A. Paczuskia and S. M. Krishnan, surgical product failure analysis.docx - 27-Paczuski-Surgical-Robotics-paper.pdf. .

[18]K. Aho, “The robotics industry: creating jobs, closing the skills gap. - Free Online Library.” [Online]. Available: http://www.thefreelibrary.com/The+robotics+industry%3a+creating+jobs%2c+closing+the+skills+gap.-a0432065390. [Accessed: 10-Nov-2015].

[19]J. Echávarri, M. Ceccarelli, G. Carbone, C. Alén, J. L. Muñoz, A. Díaz, and J. M. Munoz-Guijosa, “Towards a safety index for assessing head injury potential in service robotics,” Advanced Robotics, vol. 27, no. 11, pp. 831–844, Sep. 2013.