Investigating Learning from Demonstration in Imperfect and Real World Scenarios

Erin Hedlund-Botti
erin.botti@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Matthew C. Gombolay
matthew.gombolay@cc.gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

ABSTRACT

As the world's population is aging and there are growing shortages of caregivers, research into assistive robots is increasingly important. Due to differing needs and preferences, which may change over time, end-users will need to be able to communicate their preferences to a robot. Learning from Demonstration (LfD) is one method that enables non-expert users to program robots. While a powerful tool, prior research in LfD has made assumptions that break down in real-world scenarios. In this work, we investigate how to learn from suboptimal and heterogeneous demonstrators, how users react to failure with LfD, and the feasibility of LfD with a target population of older adults.

CCS CONCEPTS

• **Human-centered computing → User studies**; **Empirical studies in HCI**; *Accessibility design and evaluation methods.*

KEYWORDS

learning from demonstration, human-robot interaction, user studies, personalization, assistive robots

ACM Reference Format:
Erin Hedlund-Botti and Matthew C. Gombolay. 2023. Investigating Learning from Demonstration in Imperfect and Real World Scenarios. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23 Companion), March 13–16, 2023, Stockholm, Sweden.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3568294.3579980

1 INTRODUCTION

Burnout among caregivers for older adults is a critical problem due to an aging population and shortages of nurses, caregivers, and space at facilities [18, 29]. One solution to enable older adults to age in place is to develop assistive in-home care robots. These robots will need to be adaptable because everyone's home environments, preferences, and needs are different and change over time. People will need to be able to communicate their preferences to robots in a way that does not require them to be expert programmers.

Learning from Demonstration (LfD) is one method for enabling non-expert users to teach robots new skills [2]. LfD is a paradigm where the human provides a demonstration of a task and the robot learns from the demonstration. While LfD has the promise of enabling non-expert users to program robots, LfD algorithms often assume that the demonstrator is an expert and will provide perfect demonstrations [21]. However, in the real world, demonstrators may provide suboptimal demonstrations (i.e., inefficient, noisy, or incorrect demonstrations [9]) due to inexperience, fatigue, or lack of understanding [1, 17, 28]. For example, older adults may have less experience with robots or a physical or cognitive limitation that makes demonstrating a task more difficult.

In addition to potentially providing suboptimal demonstrations, people are heterogeneous; they have differing preferences, knowledge, and teaching styles [23]. Prior work has investigated how to learn from suboptimal human demonstrators using active learning [10] and inverse reinforcement learning [6–8], but does not account for heterogeneity or intentional multi-modality from demonstrators. This led us to our first research question: **How can robots learn from suboptimal and heterogeneous demonstrators?** Furthermore, if a robot understands how a demonstrator is suboptimal, can the robot teach people how to be better demonstrators? Prior work has shown that human demonstrators benefit from this type of feedback or transparency from the learning system [1, 30].

Even if a demonstrator can provide perfect demonstrations, the robot may still make mistakes and fail during the learning process due to noisy or changing environments. Additionally, robots may fail due to differences in a user's understanding of the system's functionality and limits [20]. Many prior works have shown that robot failure decreases trust in the robot [11, 12, 19, 22]. However, we need to understand how people react to failure when they are part of the teaching process and may be partially responsible for the failure. In our second research question, we seek to understand: **How does robot failure during LfD impact people's perceptions of the robot and themselves?**

After examining LfD in imperfect scenarios that will occur in the real world (i.e., robot failure and human suboptimality), we aim to investigate current LfD algorithms with a target population of end-users and determine any necessary improvements to ensure LfD is feasible with real end-users. While LfD has the potential of helping older adults personalize and teach care robots, limited prior work has investigated LfD with an older adult population. Our third research question is: **How does a target population of older adults respond to teaching a robot with LfD?** Saunders et al. created an interface that included an option for users to provide demonstrations via passive observation [24]. They evaluated their interface in an exploratory study with three older adult participants. We intend to conduct a larger study with a population of older adults with mild cognitive impairment (MCI) and their caregivers. Additionally, we plan to investigate more aspects of LfD, including teaching method and impact of failure.

HRI '23 Companion, March 13–16, 2023, Stockholm, Sweden

ACM ISBN 978-1-4503-9970-8/23/03.
https://doi.org/10.1145/3568294.3579980

2 RESEARCH APPROACH

Learning from Suboptimal and Heterogeneous Demonstrators. In prior work, we developed a novel framework called Mutual Information-Driven Meta-Learning from Demonstration (MIND MELD) [26]. MIND MELD learns to map a person's suboptimal demonstrations to improved demonstrations by learning a personalized embedding that encapsulates an individual's suboptimal teaching style. We evaluated our framework with a human-subjects experiment (n=42) and found that MIND MELD was better, than baseline LfD algorithms, in terms of objective robot performance metrics and human subjective metrics, such as likeability [3], perceived intelligence [3], workload [13], and trust [16]. Our novel algorithm was the first to effectively learn from suboptimal and heterogeneous demonstrators.

Since MIND MELD learns a personalized embedding that describes how a person is suboptimal, the robot can communicate to a person how they are suboptimal. In prior work, we developed Reciprocal MIND MELD, a framework where the robot provides feedback to the human demonstrator [25]. Reciprocal MIND MELD converts the personalized embedding from MIND MELD into actionable, verbal robotic feedback on how a person's demonstration was suboptimal and communicates this feedback to the demonstrator.

We conducted a series of three human-subjects experiments (n=126) to evaluate Reciprocal MIND MELD, finding that participants' suboptimality decreased, robot performance increased, and participants rated Reciprocal MIND MELD with higher trust [16] and team fluency [15]. Through Reciprocal MIND MELD, we can successfully enable suboptimal demonstrators to be better robot teachers. Reciprocal MIND MELD is important because in real world scenarios, demonstrators may need instruction and guidance on how to accurately and effectively provide demonstrations to robots.

Robot Failure and LfD. Due to suboptimal demonstrators, encountering novel situations, or a changing environment, robots will fail during the learning process (i.e., not complete the task correctly). To investigate how people respond to robot failure when teaching a robot with LfD, we conducted a human-subjects experiment (n=48) [14]. Participants taught a robot three different tasks using LfD (pick and place, inserting a tube into a slot, and retrieving a block between obstacles). These are manipulation tasks relevant to cooking and cleaning domains (e.g., loading and unloading a dishwasher). Each participant taught the robot using one of three main LfD instruction methods: kinesthetic teaching (physically moving the robot), teleoperation (using a controller to move the robot), or passive observation (completing the task oneself while the robot takes a video) [21]. After each demonstration, participants viewed the robot's performance of the task. Unknown to the participants, the success or failure of the task was predetermined by the experimenter to control performance as an independent variable.

We first compared participant impressions across LfD instruction methods. Participants found teleoperation to be the highest workload [13]. Additionally, participants had the lowest self-confidence [3] while using kinesthetic teaching and the highest while teaching using passive observation. Passive observation may be the most accessible interface in the real-world.

We then compared success and failure and found that when the robot failed, participants trusted the robot less and themselves less and participants thought that the robot trusted the participants less. When teaching a robot with LfD, failure not only impacted participants' trust in the robot, but also participants' trust in themselves and how participants think others view them. Therefore, information and feedback from the robot may help participants to improve their demonstrations and build back their self-confidence and trust in the robot. Methods such as Reciprocal MIND MELD may be crucial in effective real-world deployments of robots to improve accuracy and trust-repair.

3 FUTURE WORK

We plan to investigate the feasibility of LfD algorithms with older adults with respect to usability and acceptance. We aim to 1) determine if older adults can successfully and are willing to teach robots with current LfD methods and 2) discover potential improvements for LfD that would enhance usability and efficacy for older adults. Thus, we propose to conduct a human-subjects experiment where older adults teach a robot using LfD. Our target population will be older adults with MCI and their caregivers, from the Cognitive Empowerment Program at Emory University, because of their unique insights into how care robots could improve their lives.

To inform our study design, we conducted two focus groups with our target population. In the first focus group, we discussed what tasks would be useful for a robot to do in the home. We found that participants expressed interest in cleaning and chore type tasks, similar to prior work [27]. As such, in our study, we plan to use a kitchen domain where the tasks will include cleaning the counter and putting away dishes. The second focus group targeted which LfD teaching interfaces participants would prefer. As also found in prior work, participants expressed interest in passive observation and were wary of kinesthetic teaching [4]. However, no prior work has conducted a study where older adults teach or interact with a robot using kinesthetic teaching. Participants may be wary of an unknown concept, but may feel differently after experiencing the interface, so we intend to have participants teach a robot using kinesthetic teaching and passive observation. Additionally, we can compare results from this population to our prior work evaluating LfD interfaces [14].

We will use both quantitative and qualitative metrics in our study. After participants teach the robot, we will measure the robot's success at the task, participants' trust [16] in the robot and themselves, participant workload [13], and usability [5]. Additionally, we plan to conduct qualitative semi-structured interviews to evaluate and understand the perceived benefits and pain points of current LfD algorithms, teaching methods, and interfaces for older adults.

Finally, we plan to determine potential improvements for LfD to enable LfD to work in the real-world with a population of older adults. We will leverage our prior work [14, 25] and investigate how feedback systems can mitigate failure and improve usability, understanding, and performance for older adults using LfD.

ACKNOWLEDGMENTS

This work was supported by a NASA Early Career Fellowship (grant 80HQTR19NOA01-19ECF-B1), a gift from Konica Minolta, and the National Science Foundation (grants 1545287, 20-604, and IIS-2112633).

REFERENCES

[1] Saleema Amershi, Maya Cakmak, W. Bradley Knox, and Todd Kulesza. 2014. Power to the people: The role of humans in interactive machine learning. *AI Magazine* 35, 4 (2014), 105–120. https://doi.org/10.1609/aimag.v35i4.2513

[2] Brenna Argall, Sonia Chernova, Manuela M. Veloso, and Brett Browning. 2009. A survey of robot learning from demonstration. *Robotics and Autonomous Systems* 57, 5 (May 2009), 469–483. https://doi.org/10.1016/j.robot.2008.10.024

[3] Christoph Bartneck, Elizabeth Croft, and Dana Kulic. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics* 1, 1 (2009), 71–81. https://doi.org/10.1007/s12369-008-0001-3

[4] Jenay M. Beer, Akanksha Prakash, Cory-Ann Smarr, Tracy L. Mitzner, Charles C. Kemp, and Wendy A. Rogers. 2012. "Commanding Your Robot" Older Adults' Preferences for Methods of Robot Control. *Proceedings of the Human Factors and Ergonomics Society ... Annual Meeting. Human Factors and Ergonomics Society. Annual meeting* 56, 1 (Sept. 2012), 1263–1267. https://doi.org/10.1177/1071181312561224

[5] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (11 1995).

[6] Daniel Brown, Wonjoon Goo, Prabhat Nagarajan, and Scott Niekum. 2019. Extrapolating Beyond Suboptimal Demonstrations via Inverse Reinforcement Learning from Observations. In *Proceedings of the 36th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 97)*, Kamalika Chaudhuri and Ruslan Salakhutdinov (Eds.). PMLR, 783–792.

[7] Daniel S. Brown, Wonjoon Goo, and Scott Niekum. 2020. Better-than-Demonstrator Imitation Learning via Automatically-Ranked Demonstrations. In *Proceedings of the Conference on Robot Learning (Proceedings of Machine Learning Research, Vol. 100)*, Leslie Pack Kaelbling, Danica Kragic, and Komei Sugiura (Eds.). PMLR, 330–359.

[8] Letian Chen, Rohan Paleja, and Matthew Gombolay. 2021. Learning from Suboptimal Demonstration via Self-Supervised Reward Regression. In *Proceedings of the 2020 Conference on Robot Learning (Proceedings of Machine Learning Research, Vol. 155)*, Jens Kober, Fabio Ramos, and Claire Tomlin (Eds.). PMLR, 1262–1277.

[9] Sonia Chernova and Andrea L. Thomaz. 2014. *Robot Learning from Human Teachers*. Morgan & Claypool Publishers.

[10] Sonia Chernova and Manuela M. Veloso. 2009. Interactive Policy Learning through Confidence-Based Autonomy. *Journal of Artificial Intelligence Research* 34 (2009), 1–25.

[11] Joseph DelPreto, Jeffrey I. Lipton, Lindsay Sanneman, Aidan J. Fay, Christopher Fourie, Changhyun Choi, and Daniela Rus. 2020. Helping Robots Learn: A Human-Robot Master-Apprentice Model Using Demonstrations via Virtual Reality Teleoperation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 10226–10233. https://doi.org/10.1109/ICRA40945.2020.9196754

[12] Peter A. Hancock, Deborah R. Billings, Kristin E. Schaefer, Jessie Y.C. Chen, Ewart J. De Visser, and Raja Parasuraman. 2011. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Human Factors* 53, 5 (2011), 517–527. https://doi.org/10.1177/0018720811417254 ISBN: 0018720811417.

[13] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload* 1, 3 (1988), 139–183.

[14] Erin Hedlund, Michael Johnson, and Matthew Gombolay. 2021. The Effects of a Robot's Performance on Human Teachers for Learning from Demonstration Tasks. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21)*. Association for Computing Machinery, New York, NY, USA, 207–215. https://doi.org/10.1145/3434073.3444664 event-place: Boulder, CO, USA.

[15] Guy Hoffman. 2019. Evaluating Fluency in Human-Robot Collaboration. *IEEE Transactions on Human-Machine Systems* 49, 3 (2019), 209–218. https://doi.org/10.1109/THMS.2019.2904558

[16] Jiun-Yin Jian, Ann Bisantz, and Colin Drury. 2000. Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics* 4 (2000), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04

[17] Michael Laskey, Caleb Chuck, Jonathan Lee, Jeffrey Mahler, Sanjay Krishnan, Kevin Jamieson, Anca Dragan, and Ken Goldberg. 2017. Comparing human-centric and robot-centric sampling for robot deep learning from demonstrations. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*. 358–365. https://doi.org/10.1109/ICRA.2017.7989046

[18] Meredith Mealer, Ellen L. Burnham, Colleen J. Goode, Barbara Rothbaum, and Marc Moss. 2009. The prevalence and impact of post traumatic stress disorder and burnout syndrome in nurses. *Depression and Anxiety* 26, 12 (2009), 1118–1126. https://doi.org/10.1002/da.20631

[19] Neville Moray, Toshiyuki Inagaki, and Makoto Itoh. 2000. Adaptive Automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology: Applied* 6, 1 (2000), 44–58. https://doi.org/10.1037/0278-7393.6.1.44

[20] Donald A. Norman and Stephen W. Draper. 1986. *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc., USA.

[21] Harish Ravichandar, Athanasios S. Polydoros, Sonia Chernova, and Aude Billard. 2020. Recent Advances in Robot Learning from Demonstration. *Annual Review of Control, Robotics, and Autonomous Systems* 3, 1 (May 2020), 297–330. https://doi.org/10.1146/annurev-control-100819-063206 Publisher: Annual Reviews.

[22] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. Would You Trust a (Faulty) Robot?: Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. *ACM/IEEE International Conference on Human-Robot Interaction* 2015-March (2015), 141–148. https://doi.org/10.1145/2696454.2696497 ISBN: 9781450328821 Publisher: ACM.

[23] Claude Sammut. 1992. Automatically Constructing Control Systems by Observing Human Behaviour. *Second International Inductive Logic Programming Workshop* (1992).

[24] Joe Saunders, Dag Sverre Syrdal, Kheng Lee Koay, Nathan Burke, and Kerstin Dautenhahn. 2016. "Teach Me–Show Me"—End-User Personalization of a Smart Home and Companion Robot. *IEEE Transactions on Human-Machine Systems* 46, 1 (2016), 27–40. https://doi.org/10.1109/THMS.2015.2445105

[25] Mariah L. Schrum, Erin Hedlund-Botti, and Matthew Gombolay. 2022. Reciprocal MIND MELD: Improving Learning From Demonstration via Personalized, Reciprocal Teaching. In *6th Annual Conference on Robot Learning*.

[26] Mariah L. Schrum, Erin Hedlund-Botti, Nina Moorman, and Matthew C. Gombolay. 2022. MIND MELD: Personalized Meta-Learning for Robot-Centric Imitation Learning. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI '22)*. IEEE Press, 157–165.

[27] Cory-Ann Smarr, Tracy L. Mitzner, Jenay M. Beer, Akanksha Prakash, Tiffany L. Chen, Charles C. Kemp, and Wendy A. Rogers. 2014. Domestic Robots for Older Adults: Attitudes, Preferences, and Potential. *International journal of social robotics* 6, 2 (April 2014), 229–247. https://doi.org/10.1007/s12369-013-0220-0

[28] Jonathan Spencer, Sanjiban Choudhury, Matthew Barnes, Matthew Schmittle, Mung Chiang, Peter Ramadge, and Siddhartha Srinivasa. 2020. Learning from Interventions: Human-robot interaction as both explicit and implicit feedback. In *Robotics: Science and Systems*, Marc Toussaint, Antonio Bicchi, and Tucker Hermans (Eds.). MIT Press Journals, United States. https://doi.org/10.15607/RSS.2020.XVI.055

[29] Adriana Tapus, Maja J. Mataric, and Brian Scassellati. 2007. Socially assistive robotics [Grand Challenges of Robotics]. *IEEE Robotics & Automation Magazine* 14, 1 (March 2007), 35–42. https://doi.org/10.1109/MRA.2007.339605 Conference Name: IEEE Robotics & Automation Magazine.

[30] Andrea L. Thomaz and Cynthia Breazeal. 2008. Teachable Robots: Understanding Human Teaching Behavior to Build More Effective Robot Learners. *Artif. Intell.* 172, 6-7 (April 2008), 716–737. https://doi.org/10.1016/j.artint.2007.09.009