Training Human Teacher to Improve Robot Learning from Demonstration: A Pilot Study on Kinesthetic Teaching

Conference Paper · November 2020

DOI: 10.1109/RO-MAN47096.2020.9223430

4 authors, including:

Maram Sakr
University of British Columbia - Vancouver

Elizabeth A. Croft
Monash University (Australia)

2020 29th IEEE International Conference on
Robot and Human Interactive Communication (RO-MAN)
August 31 - September 4, 2020 - Virtual Conference

Training Human Teacher to Improve Robot Learning from Demonstration: A Pilot Study on Kinesthetic Teaching

Maram Sakr*[1,2], Martin Freeman[3], H.F. Machiel Van der Loos[1], and Elizabeth Croft[2]

***Abstract*— Robot Learning from Demonstration (LfD) allows robots to implement autonomous manipulation by observing the movements executed by a demonstrator. As such, LfD has been established as a key element for useful user interactions in everyday environments. Kinesthetic teaching, a teaching technique within LfD, entails physically guiding the robot to achieve a task. When demonstrating complex actions on a multi-DoF manipulator, novice users typically encounter difficulties with trajectory continuity and joint orientation, necessitating training by an expert. A comparison between different training approaches is conducted in a study of nine novice users. These approaches are kinesthetic, observational and discovery-learning. The kinesthetic method utilizes record and playback functions implemented on a 7-DoF Barrett Technology WAM robot. A novice user passively holds the arm while an expert's trajectory is replayed. A visual demonstration by the expert is used for the observational training group. The discovery-learning group does not receive an expert demonstration; they use trial-and-error to produce the trajectory on their own. Task-space performance is evaluated pre- and post-training for each user to determine the relative and absolute performance improvements of the groups across the three training approaches. Absolute performance improvements are compared to the performance of an expert and a minimum-jerk trajectory to gauge how skillful the participant becomes with respect to the expert. The kinesthetic approach shows superior indicators of performance in trajectory similarity to the minimum-jerk trajectory with 39% and 13% improvement over the observational and discovery methods, respectively. Observational training shows greater improvement in terms of the smoothness of the velocity profile with 32.7% compared to 29.5% and 21.9% for both discovery and kinesthetic training, respectively.**

I. INTRODUCTION

Over the past decade, automation technology has continued to advance the production industry by phasing out the human-centricity of complex assembly tasks and increase the use of more cost-efficient and accurate robotic manipulators [1]. Such robots are becoming more ubiquitous in everyday settings, and have begun to permeate into educational, medical and social environments [2]. Traditional programming methods for such manipulators are typically outside the skill-set of the average user [3] and thus alternative forms of task programming have been developed.

Robot Learning from Demonstration (LfD) is the field of research that minimizes or eliminates the "technical" programming requirement when working with robots. LfD

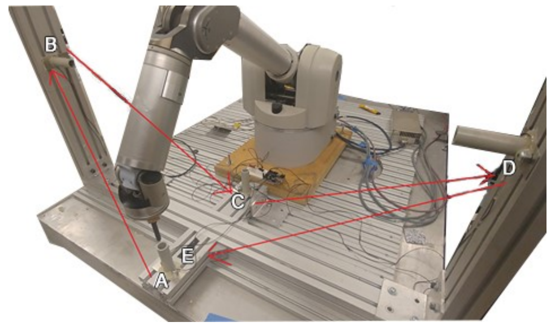

Fig. 1: Task setup with the WAM robot. Letters from A to E represent the sequence for pressing buttons in the pipes.

can be defined as the paradigm that allows robots to perform tasks after observing a teacher performing them [4]. LfD opens the doors for domain-experts without robotics or programming background to teach the robot many tasks. Louie et al., [5] use the LfD to teach a social robot how to facilitate a group activity for older adults. Murali et al., [6] automate the performance of surgical subtasks through the surgeon's demonstration. In addition, LfD has been used to teach the robots many household tasks like wiping a surface [7], peeling vegetables [8], and dish-washing [9].

Kinesthetic teaching is one of the most commonly used interfaces for teaching robots by demonstration. It allows the demonstrator to physically hold the robot and guide it to do a task. Figueroa et al. [10] used this interface to teach a robot to roll pizza dough. Others have used kinesthetic teaching in more complex tasks such as ball paddling [11], box flipping using chopsticks [12] and knot tying [13]. Kinesthetic teaching has several advantages for LfD as follows: i) there is no correspondence problem as the teacher directly guides the robot, ii) demonstrations are restricted to the kinematic limits (e.g., joint limits, workspace) of the robot, and iii) there is no need for extra instrumentation beyond the robot's own sensors and actuators. While kinesthetic teaching has been extensively used in the LfD literature, it may be challenging for novice users who don't have experience with robots. Typically, it is the student/researcher who teaches the robot in the majority of the LfD literature. In practical LfD applications, the teacher will not be an expert in robotics and hence it is important to ensure that the user is able to teach the robot beforehand. This is important because the robot's task performance is bounded by how well the user can smoothly and efficiently perform the task as noted in [4].

In recent literature, researchers have tackled this issue from different perspectives. Grollman and Billard [14] have

[1]Collaborative Advanced Robotics and Intelligent Systems (CARIS) Laboratory, Mechanical Engineering, University of British Columbia.
[2]Electrical and Computer Systems Engineering, Monash University
[3]Stanford University
*Email: maram.sakr@ubc.ca

explored the possibility of using failed demonstration in robot learning. They consider these failed demonstrations as an example of ”what not to do”. The main problem with this approach is that they define the failed demonstrations as ones which did not achieve the task goal. This definition overlooks the demonstrations that achieve the task goal but with many unnecessary movements that affects the robot learning [4]. Others proposed the use of keyframe demonstration [15] at which the user will ask the robot to record only around selected waypoints and then they interpolate between them. This approach helps eliminate the jerky data in between the waypoints. The issue with this approach is that the demonstration data lake the time information and some users struggle to define the critical waypoints of the task and find it a mentally-demanding task. Cakmak and Takayama [16] have explored the effect of instructional materials design to teach people how to teach robots. They found that most of the participants tend to work with the robot without completing the provided instructions which results in poor demonstrations. Also, they found that using video instruction is better than the typical written user manual. Recently, Sena and Howard [17] propose a method for quantifying teaching behaviour of the human demonstrator. They show that a significant improvement of robot performance was achieved by providing a feedback to the teacher after the task demonstration. They target the problems of undemonstrated states, ambiguous demonstration, and incorrect demonstrations. While these problems affect the robot's skill learning, this approach overlooks the question of whether the human teacher is able to demonstrate the task clearly to the robot.

While researchers study the use of failed demonstrations, provide different demonstration techniques or provide feedback to the human teachers, others explore different teaching interfaces that are easy for end-users and that bring clear data for the robot. Liang et al., [18] proposed using the Intuitive Surgical da Vinci as an interface to teleoperate an ABB YuMi robot. This facilitates the collection of time-synchronized vision and motion data. Also, it allows the user to teach the robot some complex bi-manual tasks. This approach provides guidelines for teaching interface design, but using the da Vinci's controllers is not always intuitive for novice users [19], [20].

In this paper, we provide a training approach for novice users to kinesthetically teach the robot. The paper contributions are as follows:

1) Proposing a training approach for the user before teaching the robot by demonstration.
2) Exploring different training techniques to find the best suitable one.
3) Proposing some performance metrics for evaluating the demonstration data.

II. PROPOSED SYSTEM

A. System Overview

We use the 7-DOF Barrett Whole Arm Manipulator (WAM) [21] with a task-customized end-effector as our training platform. In addition, a ROS repository that makes use of the functionality of the Barrett WAM C++ library is used. A python script is used for recording the demonstration data of a task kinesthetically taught by an expert (one of the paper's authors) to the WAM robot. These demonstration data will be played back on the robot while a novice user holds it passively to feel the robot motions. This approach called *hand-over-hand* or *kinesthetic* training approach. Several studies [22], [23] support that this hand-over-hand teaching may be an effective way to transfer complex skills from the expert to the novice. The haptic input in the kinesthetic teaching has shown to be effective in both the sports and surgical worlds [24], [25], [23] to teach motor skills, and is thus examined in this study as a viable form of teaching novice demonstrators.

B. Training Groups

In this paper, we propose two training techniques that can be used to train a novice user how to kinesthetically teach a robot. In order to realize the benefits of the proposed two techniques, we compare them to the *discovery group* at which the user will discover via trial and error how to teach the robot without any guidance. Thus, we have a total of three training groups as follows:

1) **Discovery Group**: this group will learn by exploring. The trainees will be given some time to practice with the robot and self-reflect on their performance and determine an area they can apply effort in to improve in their next demonstrations.
2) **Observational Group**: this group will learn by watching an expert teaching the robot a specific task. They generally focus their observation on the kinematic motion of the WAM as a result of expert's movements.
3) **Kinesthetic Group**: this group will learn by doing. Trainees will hold the robot passively while it plays back a previously taught trajectory by an expert to accomplish a task.

C. Robotic Task

It is desired to identify a task that validly differentiates expert and novice users. A series of criterion to describe the elements that make up a complex task are identified as follows:

- Manipulation of end-effector in 3D space
- Orthogonal changes in end-effector orientation
- Requirement of moderate positional accuracy

A fundamental task was constructed to incorporate the range of complexity elements above. The task is designed to be performance-oriented rather than goal-oriented, and as such each participant will be familiarized enough with the arm to complete the task in full. We chose a modified version of a commonly-used peg-in-hole task. In this task, the participant will move the arm and orient the end-effector to sequentially press the orthogonally-oriented buttons located within a pipe that forces the end-effector to adopt a normal orientation, as shown in Fig.1. This task actively incorporates the participant's ability to implicitly manipulate the end-effector in a 3D space through direct, simultaneous manipulation of two to four joints, as well as the the skillful orienting of the end-effector in a variety of ways due to the depressed and orthogonal positioning of the buttons.

Participants will follow the demonstration route depicted in Fig.1, pressing the buttons at each of the checkpoints along the way. The push buttons are located 4cm down a PVC tube and require the robot's end-effector be located normal to the checkpoint plane. Participants are not restricted to the trajectory or route they take between any checkpoints. When the button is pressed, the instrumentation setup plays a notification sound so that the user does not need to look away from the task.

III. EXPERIMENT

We conducted a study to evaluate the effectiveness of our proposed training framework for robot kinesthetic teaching. We recruited nine study participants from University of British Columbia who had no prior experience with robots. Novice users, especially ones who have never manipulated a robot, will generally exhibit a certain naivety during their very first interaction events. Thus, we start the study with allowing the participants to familiarize themselves with the robot for three minutes. Obtaining references during this familiarization phase for the feel, weight, responsiveness, range of motion and other capabilities of the arm as well as moving past the initial shock factor of first manipulating a robot will help standardize the user pool prior to the task trials. During this time, participants are instructed to actuate all of the joints and familiarize themselves with the control of the overall motion of the arm from the hand position designated by a colored tape to avoid pinch-points and to aid the standardization of the user's demonstrations.

After the familiarization phase, all participants undergo their *baseline evaluations* for the task in the benchmark phase. Instructions are given to the participant to emphasize demonstration of a continuous, smooth and efficient trajectory that they might expect from a computerized program, and that each checkpoint should be reached as a continuous node of the overall trajectory rather than a series of individual trajectories between checkpoints. Participants are also informed that the maximum time to complete the task is 60 seconds, and if they cannot complete the task within the time, analysis of their partial trajectory can be performed with a non-completion penalty.

Following this, the participants are divided into three training groups: discovery, observational and kinesthetic. A total of five training trials will be conducted where each trial consists of the *"teaching"* and *"application"* phases of their group's appropriate condition as follows:

- **Kinesthetic group**: for the teaching phase, the participants are told to position their hands and feet on marked locations with the colored tape, gripping the WAM's links as if they were preparing to demonstrate their own trajectory. Then, they are instructed to direct their eyes toward only the next immediate waypoint during the playback. They are also instructed to mentally focus on how the joints in their own arms and hand must move to accommodate the playback trajectory. Once the participant has confirmed they are ready, an expert trajectory will be played on the robot as the participant performs the attentive procedure instructed above. Immediately after playback of the trajectory used in the teaching phase, the participant will relinquish control of the arm as it is moved back to the start position. The participant is told to attempt their best rendition of the expert trajectory using what they have felt from the previous teaching trial.
- **Observational group**: for the teaching phase, the participant is told to stand at a location indicated by colored tape. This location is different from the Kinesthetic condition, as the participant in the observation condition should observe the expert demonstration from an orientation similar to the expert's. The expert will take their place and instruct the participant to generally focus their observation on the kinematic motion of the WAM as a result of expert's movements. Once the expert completes demonstration of the task, the participant will configure themselves in the same orientation as the expert, placing their hands and feet at the appropriate locations marked with colored tape. The participant is told to attempt their best rendition of the expert trajectory using what they have observed from the previous teaching trial.
- **Discovery group**: participants do not receive a formal teaching phase, but are told to take 30 seconds between trials to self-reflect on their previous demonstration and determine an area they can apply effort in to improve in their next demonstration. After the self-reflection period, the participants will place their hands at the appropriate locations marked with colored tape. A reminder will be given to the participant to try to best apply the element of improvement they came up within the self-reflection period.

After the training, all participants are informed that their previous trials were for training purposes, and that their performance for those trials are not evaluated. Instead, they are told that there is one final *evaluation* trial. During this evaluation trial, participants are instructed to emphasize demonstration of a continuous, smooth and efficient trajectory in the shortest time as possible.

IV. EVALUATION

In order to evaluate the resulting trajectory from each trial, we compare it with a minimum-jerk trajectory as well as the expert's trajectory. The minimum-jerk principle has been used to predict both qualitative and quantitative features of human arm movements, whereby suggesting that the trajectory that minimizes the integral of squared jerk very closely resembles the naturally smooth trajectories our limbs take [26]. Previous work [27], [28] suggest that a constrained minimum-jerk model can appropriately predict the speed profiles of complex, constrained arm movements.

To generate this minimum-jerk trajectory, a MATLAB script is used to take the constrains points as an input and generate the joint states at each point. These constrains

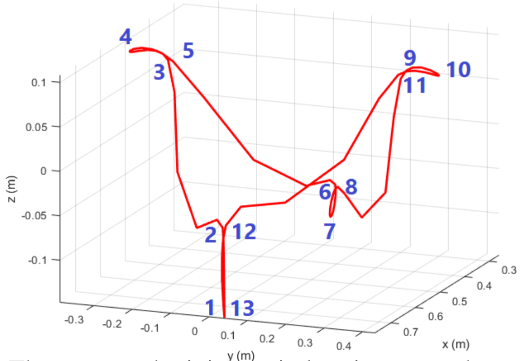

Fig. 2: The generated minimum-jerk trajectory and waypoints sequence

points represent the start and end positions of the task, as well as intermediate positions outside and inside each pipe. We use polynomial splines to interpolate a number of segments between each waypoint in the trajectory such that a maximum smoothness is achieved when jerk is minimized as follows:

$$min \int_0^{t_f} \left| \frac{d^3x}{dt^3} \right|^2 dt \quad (1)$$

Forward kinematics for these waypoints was used to generate the Cartesian trajectory. Each waypoint in the minimum-jerk trajectory was scaled to the length of the user's trajectory time vector. Thus, we assign a time proportionality to individual actions in the minimum-jerk trajectory.

The trajectory segment relevant to insertion of the end-effector into the pipe uses three outside-inside-outside waypoints to represent the fact that real demonstration trajectories must physically remove the end-effector from the pipe before moving to the next checkpoint. The minimum-jerk trajectory is generated using an 8^{th} order interpolating polynomial in each segment. The minimum-jerk generated for comparison in our analysis, along with the sequencing of waypoints are displayed in Fig. 2.

A. Performance Metrics

Four metrics are used to analyze the task-space performance of users during their benchmarking and evaluation trials as follows:

1) *Completion time, t*: The start to finish time taken to move the end-effector through all requirements of the task.
2) *Pathlength, S*: The start to finish pathlength of the task-space trajectory. Given a task trajectory data set of n points with coordinates x_i, y_i, z_i indicating the i^{th} point, the total pathlength is calculated as follows:

$$S = \sum_{i=1}^{n-1} \sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2 + (z_{i+1}-z_i)^2} \quad (2)$$

3) *Smoothness*: Two quantitative metrics are used here to represent trajectory smoothness based on two common methods in the field as follows:

a) *Distance-based similarity to minimum-jerk trajectory, d_{mj}*: A dynamic time warping (DTW) algorithm is used to judge the similarity of the user's trajectory to the minimum-jerk one. DTW is irrespective of time of completion and does not change as either trajectory is scaled temporally, though variation in this metric can occur as the user trajectory starts to lag or lead the position minimum-jerk trajectory given equal time-domain vectors [29]. As both our minimum-jerk trajectory Q and demonstrated trajectory C are multi-dimensional time series', we must choose one of the multi-dimensional generalization forms of DTW [30]. The independent form, DTW_I summates the distances of the two trajectory's x, y, z dimensions. The dependent form, DTW_D forces all dimensions to warp identically. As our x, y, z dimensions are dependently sequences to represent an overall Cartesian position, we use the equations for DTW_D for our quantification as follows:

$$DTW_D(Q,C) = DTW(\{Q_x, Q_y, Q_z\}, \{C_x, C_y, C_z\}) \quad (3)$$

where the distance $d(q_i, c_j)$ is defined as the cumulative squared Euclidean distances of three z-normalized data points as follows:

$$d(q_i, c_j) = \sum_{m=1}^{m=3} (q_{i,m} - c_{j,m})^2 \quad (4)$$

b) *Steadiness, N_a*: The number of velocity peaks per meter has been frequently used in movement smoothness literature as a reasonable standard of smoothness [31]. It is a measure that counts the number of maxima in a given velocity profile $v(t)$ as shown in (5). Both the smoothed user trajectory, C, and minimum-jerk trajectory, Q, are differentiated to obtain the velocity series for which the peaks are counted and divided by the respective pathlength. For trajectory C, the velocity data x, y, z is smoothed with a robust quadratic regression filter to remove noise and help delineate the location of velocity peaks without losing details of the trajectory. Fig. 3 show the result of the filter window size being scaled accordingly until all notable peaks are represented. This filtering method reduces the appearance of multi-headed peak clusters, nullifying the need for strict usage of a minimum peak separation threshold. Due to the large variation in the number of small peaks that can come with different times of completion, a larger minimum peak separation threshold of 50 samples, or 0.2 seconds, is used. As a baseline for what would be considered the optimal number of peaks per

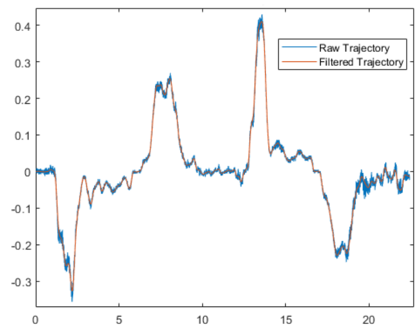

Fig. 3: Robust quadratic regression filtering of velocity data.

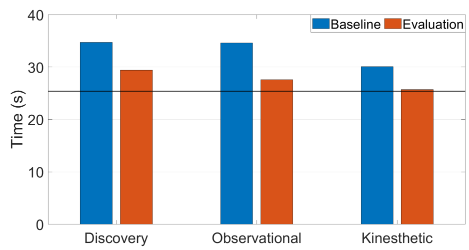

Fig. 4: Total completion time for the three training groups compared to expert's trajectory (solid line).

meter, the minimum-jerk trajectory is generated with the same time of completion as the user's trajectory and the number of peaks per meter are measured.

$$NP = \frac{|\{v(t), \frac{dv(t)}{dt} = 0 \; and \; \frac{d^2v(t)}{dt^2} < 0\}|}{S} \quad (5)$$

where $|.|$ represents the cardinality of a set and S is the traveled pathlength.

V. RESULTS

A. Completion Time

Fig. 4 shows the total completion time of the trajectory for both the baseline and evaluation trials of the novice participant's trajectory, as well as the expert trajectory. The observational group displayed the greatest improvement in time of completion at 25.2%, with the discovery and kinesthetic groups scoring much lower at 17.9% and 17.3% improvement. The kinesthetic group produced the lowest mean time of completion for their evaluation trials at 25.7s, followed closely by the observational group at 27.6s and the discovery group at 29.4s. The expert trajectory produced a time to completion of 25.4s.

B. Pathlength

Fig. 5 shows the total pathlength of the trajectory for the baseline and evaluation trials of the novice participant's trajectory, as well as the expert trajectory and minimum-jerk trajectory. The kinesthetic group exhibited the largest performance increase at 10.4%, followed by the observational group at 8% and the discovery group at 5.7%. The kinesthetic group produced the greatest absolute mean performance in

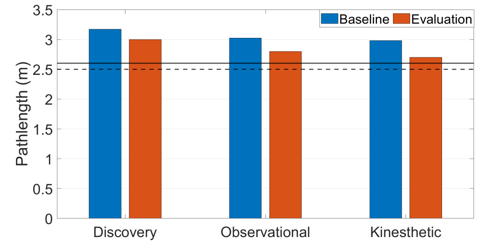

Fig. 5: Pathlength for the three training groups compared to the expert's trajectory (solid line) and the minimum-jerk trajectory (dashed line).

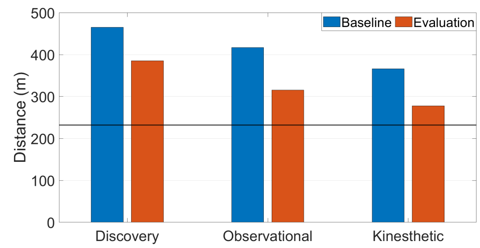

Fig. 6: Distance similarity to the minimum-jerk trajectory for the three training groups as well as the expert's trajectory (solid line).

their evaluation trials at 2.7m, followed by the observation group at 2.8m, and the discovery group at 3m. The expert produced a pathlength of 2.6m, slightly lower than all three groups, and very close to the minimum-jerk pathlength of 2.5m.

C. Distance similarity to minimum-jerk trajectory

Fig. 6 shows the mean distance similarity between the novice participants trajectory and the minimum-jerk trajectory for the baseline and evaluation trials. The distance similarity between the expert trajectory and minimum-jerk trajectory was also included. The Observational and kinesthetic groups produced superior performance increases in distance similarity at 32.2% and 31.9% respectively. The discovery group followed far behind at a 20.8% performance increase. The kinesthetic group exhibited significantly higher absolute mean performance in their final evaluation trials with a similarity of 277.4m, compared to the observational and discovery groups at 315.3m and 384.97m, respectively. The individual trials closest in performance to the expert's similarity of 231.98m were two participants in the kinesthetic group at final evaluation similarities of 246.7m and 249.6m, respectively.

D. Steadiness

Fig.7 shows the mean velocity peaks per meter detected in the filtered velocity profile of each novice participant's trajectory for both the baseline and evaluation trials. The observational group narrowly exhibited the largest performance increase at 32.7%, followed by the discovery and kinesthetic groups at 29.5% and 21.9% respectively. Thus, the observational and discovery groups displayed the greatest percentage change between baseline and evaluation trials.

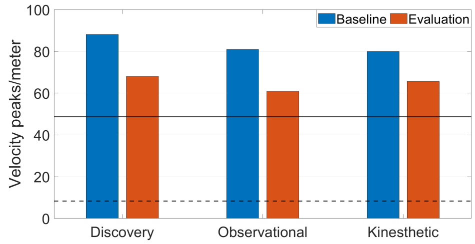

Fig. 7: Velocity peaks per meter for the three training groups compared to the expert's trajectory (solid line) and the minimum-jerk trajectory (dashed line).

The observation group also displayed the greatest absolute mean performance in their evaluation trial at 60.97 peaks per meter, followed by the kinesthetic group at 65.55 peaks per meter, and the discovery group at 68.05 peaks per meter. The expert produced a value of 48.7 peaks per meter, significantly lower than the evaluation trial means of all of the studied groups. The generated minimum-jerk trajectory produced a value of 8.3 peaks per meter, considered to be the optimal value to achieve maximum smoothness.

VI. DISCUSSION

The results from the user study show an improvement in the performance whatever the training method used which highlight the importance of the training for novice users before teaching a robot by demonstration. Observational training shows the highest improvement in both completion time and trajectory consistency. Kinesthetic training shows the highest absolute performance in both pathlength and similarity to the minimum-jerk trajectory.

A. Velocity peaks per meter

The observational and discovery groups produced better performance increases between their baseline and evaluation trials. The poor performance of the kinesthetic group may possibly be related to the micro-management of the novice to recreate the reference expert trajectory. That is, if the trajectory to perform is more explicitly defined to a participant in the kinesthetic group, then efforts to maintain that trajectory may take on a more oscillatory nature than the somewhat interpolated trajectory of the observational group. If this is the case, a greater number of practice trials, and not necessarily teaching trials, may allow the novice to adopt a more natural control of the reference trajectory they have interpreted.

B. Distance-similarity to the minimum-jerk trajectory

The observational and kinesthetic groups produced the greatest performance increase in similarity to the minimum-jerk trajectory. This was somewhat expected as we have assumed the expert trajectory is significantly guided by the minimization of jerk, and the observational and kinesthetic groups are the only groups who learn from the expert trajectory. We thus qualify the expert trajectory as an objective standard with these results as it exhibits the greatest similarity to the minimum-jerk trajectory. While the kinesthetic group started with the lowest mean baseline similarity (437m) compared to the observational and discovery baselines (470m and 502m), it also achieved both the greatest performance increase and lowest evaluation mean. This suggests that the consistent nature of the kinesthetic group's learning input may provide better performance when desiring to teach novice users constrained trajectories that require similarity. Poor discovery group performance in similarity was expected, as the participant is left to develop their own trajectory with no reference or feedback to work toward, resulting in the observation we have here of a discovery trajectory very dissimilar to both the expert and minimum-jerk trajectories.

C. Pathlength

Following what we learned from the similarity data, pathlength should have some correlation to similarity. The kinesthetic group led the performance increase in pathlength, followed by the observational and discovery groups. By the same explanation above regarding the teaching of a constrained trajectory, it follows that the kinesthetic group performs the best in pathlength as it approaches similarity of the expert trajectory. The discovery group had a mean evaluation trial of 3m, around 30cm longer than the mean of the expert trajectory. This further supports the notion that the trajectories discovered by this group can lack severely in efficiency and similarity to an ideal trajectory if not given a reference. The observational group mean was roughly 13cm longer than the kinesthetic group mean, suggesting that, despite both groups learning from a similarly efficient expert trajectory, the observational group may produce trajectories more dissimilar and inefficient than the kinesthetic group based on the difference in teaching interfaces. The novice having to perform what they observed may have encountered dissonance in their visual perception of the movement and their motor ability to implicitly control the end-effector pose as they manipulate both the wrist and elbow joints of the WAM. The learning approach in the kinesthetic group contrasts this as the novice self-inserts into the position of the expert, reducing some of the visual and physical dissonance between the motion.

D. Completion time

The observational group produced the greatest performance increase in time of completion, though this means little as this metric was emphasized to the novice to be relatively insignificant in comparison to the others. One notable thing is that the completion times of the kinesthetic group very closely approached the completion time of the expert trajectory (25.4s). If we compare this to the different mean completion times for the observational and discovery group, it may be that the kinesthetic group was able to temporally follow the expert's played back trajectory quite closely. The time each sequence in the kinesthetically-taught expert trajectory should take to complete may have been indirectly internalized by the novice more than the observational group.

VII. CONCLUSION AND FUTURE WORK

The motivation of this study is to evaluate the natural human teaching to a robot and assess the effect of the training on their demonstration data. While primitive in the scope of a more rigorous large-population analysis, this pilot study seeks to shed a light on an area of robot LfD not pervasive in the current literature. The results of this user study demonstrate that humans do not naturally provide the best possible example when they teach a robot for the first time. In addition, training and practice improve the performance whatever the technique is. Here, we are not seeking optimal performance, instead, we are seeking a higher human-like performance. This was shown by teaching the novice users using the expert's data rather than the minimum-jerk trajectory, while the latter serve as an example of an optimal trajectory. The comparison between the three training techniques shows the merits of each approach and how each training technique contributes to a specific performance metric.

We believe that the results of this paper open up several directions for future research. First, a more extensive user study with more participants is needed to validate our results. Second, integrating the three training techniques together for a multi-modal training approach to combine all merits from the three techniques. Furthermore, it would be interesting to study the skills transferability across different tasks and robotic platforms.

REFERENCES

[1] J. Markoff, "Skilled work, without the worker," *The New York Times*, vol. 18, 2012.

[2] F. Tanaka and S. Matsuzoe, "Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning," *Journal of Human-Robot Interaction*, vol. 1, no. 1, pp. 78–95, 2012.

[3] L. Major, T. Kyriacou, and O. P. Brereton, "Systematic literature review: teaching novices programming using robots," *IET software*, vol. 6, no. 6, pp. 502–513, 2012.

[4] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469–483, 2009.

[5] W. Louie, F. Despond, and G. Nejat, "Social robot learning from demonstration to facilitate a group activity for older adults," in *IEEE International Conference on Intelligent Robots and Systems Workshop on Bridging User Needs to Deployed Applications of Service Robots*, 2015, pp. 1–6.

[6] A. Murali, S. Sen, B. Kehoe, A. Garg, S. McFarland, S. Patil, W. D. Boyd, S. Lim, P. Abbeel, and K. Goldberg, "Learning by observation for surgical subtasks: Multilateral cutting of 3d viscoelastic and 2d orthotropic tissue phantoms," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 1202–1209.

[7] J. Kim, N. Cauli, P. Vicente, B. Damas, F. Cavallo, and J. Santos-Victor, ""icub, clean the table!" a robot learning from demonstration approach using deep neural networks," in *2018 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. IEEE, 2018, pp. 3–9.

[8] N. Figueroa and A. Billard, "Learning complex manipulation tasks from heterogeneous and unstructured demonstrations," in *IROS Workshop on Synergies between Learning and Interaction*, 2017.

[9] K. Okada, T. Ogura, A. Haneda, J. Fujimoto, F. Gravot, and M. Inaba, "Humanoid motion generation system on hrp2-jsk for daily life environment," in *IEEE International Conference Mechatronics and Automation, 2005*, vol. 4. IEEE, 2005, pp. 1772–1777.

[10] N. Figueroa, A. L. P. Ureche, and A. Billard, "Learning complex sequential tasks from demonstration: A pizza dough rolling case study," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Ieee, 2016, pp. 611–612.

[11] J. Kober and J. Peters, "Learning motor primitives for robotics," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 2112–2118.

[12] P. Pastor, M. Kalakrishnan, S. Chitta, E. Theodorou, and S. Schaal, "Skill learning and task outcome prediction for manipulation," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 3828–3834.

[13] J. Schulman, J. Ho, C. Lee, and P. Abbeel, "Learning from demonstrations through the use of non-rigid registration," in *Robotics Research*. Springer, 2016, pp. 339–354.

[14] D. H. Grollman and A. Billard, "Donut as i do: Learning from failed demonstrations," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 3804–3809.

[15] B. Akgun, M. Cakmak, J. W. Yoo, and A. L. Thomaz, "Trajectories and keyframes for kinesthetic teaching: A human-robot interaction perspective," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 2012, pp. 391–398.

[16] M. Cakmak and L. Takayama, "Teaching people how to teach robots: The effect of instructional materials and dialog design," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 2014, pp. 431–438.

[17] A. Sena and M. J. Howard, "Quantifying teaching behaviour in robot learning from demonstration," *arXiv preprint arXiv:1905.04218*, 2019.

[18] J. Liang, J. Mahler, M. Laskey, P. Li, and K. Goldberg, "Using dvrk teleoperation to facilitate deep learning of automation tasks for an industrial robot," in *2017 13th IEEE Conference on Automation Science and Engineering (CASE)*. IEEE, 2017, pp. 1–8.

[19] A. N. Sridhar, T. P. Briggs, J. D. Kelly, and S. Nathan, "Training in robotic surgery—an overview," *Current urology reports*, vol. 18, no. 8, p. 58, 2017.

[20] A. E. Abdelaal, P. Mathur, and S. E. Salcudean, "Robotics in vivo: A perspective on human–robot interaction in surgical robotics," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, 2020.

[21] "Barretttech wam arm," 2018, https://www.barrett.com/wam-arm/.

[22] D. Pinzon, R. Vega, Y. P. Sanchez, and B. Zheng, "Skill learning from kinesthetic feedback," *The American Journal of Surgery*, vol. 214, no. 4, pp. 721–725, 2017.

[23] A. E. Abdelaal, M. Sakr, A. Avinash, S. K. Mohammed, A. K. Bajwa, M. Sahni, S. Hor, S. Fels, and S. E. Salcudean, "Play me back: A unified training platform for robotic and laparoscopic surgery," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 554–561, 2018.

[24] T. Yang, J. Liu, W. Huang, Y. Su, L. Yang, C. K. Chui, M. H. Ang, and S. K. Chang, "Mechanism of a learning robot manipulator for laparoscopic surgical training," in *Intelligent Autonomous Systems 12*. Springer, 2013, pp. 17–26.

[25] J. L. Patton and F. A. Mussa-Ivaldi, "Robot-assisted adaptive training: custom force fields for teaching movement patterns," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 4, pp. 636–646, 2004.

[26] N. Hogan, "An organizing principle for a class of voluntary movements," *Journal of Neuroscience*, vol. 4, no. 11, pp. 2745–2754, 1984.

[27] E. Todorov and M. I. Jordan, "Smoothness maximization along a predefined path accurately predicts the speed profiles of complex arm movements," *Journal of Neurophysiology*, vol. 80, no. 2, pp. 696–714, 1998.

[28] F. Amirabdollahian, R. Loureiro, and W. Harwin, "Minimum jerk trajectory control for rehabilitation and haptic applications," in *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, vol. 4. IEEE, 2002, pp. 3380–3385.

[29] A. Mueen and E. Keogh, "Extracting optimal performance from dynamic time warping," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 2129–2130.

[30] M. Shokoohi-Yekta, B. Hu, H. Jin, J. Wang, and E. Keogh, "Generalizing dynamic time warping to the multi-dimensional case requires an adaptive approach (sdm 2015)," in *2015 SIAM International Conference on Data Mining.-http://www. cs. ucr. edu/~ eamonn/Multi-Dimensional_DTW_Journal. pdf (last access: 18.12. 2015)*.

[31] S. Balasubramanian, A. Melendez-Calderon, A. Roby-Brami, and E. Burdet, "On the analysis of movement smoothness," *Journal of neuroengineering and rehabilitation*, vol. 12, no. 1, p. 112, 2015.