Article | Received 17 December 2025; Revised 10 February 2026; Accepted 15 February 2026; Published 23 March 2026
https://doi.org/10.55092/rl20260006

Automatic estimation and evaluation of multi-objective human preferences for Learning from Demonstration

Brendan Hertel[1],†,*, Tam Nguyen[2],†, Maria Eugenia Cabrera[2] and Reza Azadeh[1]

[1] Persistent Autonomy and Robot Learning (PeARL) Lab, University of Massachusetts Lowell, Lowell, USA

[2] Assistive Robots and Accessibility (ARA) Lab, University of Massachusetts Lowell, Lowell, USA

† These authors contributed equally to this work.

* Corresponding author; E-mail: brendan_hertel@student.uml.edu.

Highlights:

- We design and test an interface where users can select their preference of robot execution and automatically learns user preference over time, allowing automatic preferential execution.
- We develop a novel Learning from Demonstration method which can generate several optimal solutions aligning with users' unique preferences.
- We collect information on user preferences in several user studies, evaluating preferences between tasks and users.

Abstract: When robots are in the hands of end-users, they should perform tasks according to the preferences of those users. However, it is currently impractical to discover the preferences from those end-users without user-specific and task-specific feedback. To remedy this, we design and test an interface which can learn user preferences over time, eventually predicting their preference for the execution of any given task. Additionally, we collect data on user preferences from these interactions, aiming to find an overall preference of execution across tasks and across users. We validate several assumptions on user preference in a study ($N = 10$) which presents limited preference options to users. This validates our use of Pareto front options and choice of objectives in optimization. With the knowledge from this small study, we use an interface to let users explore their preference space for several unique tasks in a larger user study ($N = 20$). Our contributions include: (a) a novel Learning from Demonstration formulation for generating preference-aligned skill reproductions via multi-objective optimization, (b) a Graphical User Interface (GUI) for human-in-the-loop preference discovery, and (c) an analysis of inferred preferences across users and tasks. Through this analysis we find that individual users and tasks have specific preferences, but certain trends can be highlighted, like the desire for smooth and optimal robot execution.

Keywords: human-robot interaction; Learning from Demonstration; preference learning; human-in-the-loop interaction; multi-objective optimization

1. Introduction

As robots become increasingly integrated into human environments, it is equally essential for humans to understand robot motion and for robots to move in ways that align with human preferences. Extensive research has been conducted on the design of robots to make them more human-like and acceptable [1]. In addition, much effort has been made to enable robots to move in more human-like ways, both in manipulation [2] and locomotion [3]. While human-like motion is often thought to enhance the social acceptability of robots [4], it may not necessarily reflect the most effective or preferred way for robots to move around humans.

Often, human preference is used to inform reward functions which are then learned using reinforcement learning (RL) to create robot motion [5] or behavior [6]. However, RL can be difficult to optimize, and often preference-based methods present a multitude of choices for the user to make or rank. When it comes to learning robot motion, Learning from Demonstration (LfD) [7] can represent many features through a single demonstration from a user. However, in LfD, robots are only able to use the information provided in the demonstrations, which may be suboptimal, biased, or otherwise flawed [8–10]. There may be a preferred way for the robot to reproduce a task which is different than what was demonstrated.

For a human to trust a robot, it must successfully be able to convey its intent to the human somehow [11,12]. This can be achieved by making the robot's movement predictable and unsurprising, or moving in such a way that the robot is able to clearly convey its intent and goal [11]. The legibility and predictability of a robot's actions play a great role in Human-Robot Interaction (HRI). Knowing what a user expects out of the robot allows for the implementation and tuning of a robot's movements to align with those expectations. User's previous experiences and exposure to robots also affect an individual's interpretation of predictability of robot motion [13]. Allowing users to correct a robot when the user deems its movements to be unpredictable also allows a better understanding of robot legibility and predictability, providing opportunity for learning user preference [14]. With corrections, when and how users correct a robot's movement serves as valuable information on when and how an action can be improved upon in the future. Corrections, however, are a reactive method for creating human-preferred robot motions. Alternatively, the robot can preemptively query the user, in a framework known as active learning [15], which we adopt in this work.

Recently, there has been interest in understanding human preference for robot motion or actions [5]. In these preference-based learning methods, a human, often in-the-loop, is asked to rank or select trajectories based on their preference. In [6], users customize robot behaviors through exploratory search, leading to a set of user-aligned features. However, in some cases these user-aligned features are no better than randomly selected features. Some methods automate this process without involving a user [16], generating noisier variations of trajectories which are assumed to be less preferred. Other methods take this assumption even further, assuming that the user is limited in their ability to demonstrate, and

reproductions should also be simpler executions [17]. In [18], a collaborative human-robot performance cost is found based on human behavior in a prosthetic walking scenario. Unlike these methods, which either query the user for their preference or assume it for them, our method combines both, assuming the user preference only after it has been learned over time through queries, leading to user-specific preference learning and estimation.

Works that attempt to learn across different tasks are known as transfer learning methods [19]. In transfer learning, the information learned in one task or environment is transferred to a new one, often with some retraining. One of the most common cases for transfer learning is simulation to real-world (sim2real) transfer [20]. This allows for extensive training of a task in a no-risk simulation environment before executing potentially dangerous tasks in the real-world. Some works learn parameters for one task which can be transferred to a new one to speed up training [21]. To the best of our knowledge, no other work has attempted to transfer preferences between robotics tasks, although there is work in preference transfer for other domains [22].

Often, to understand user preferences, a Graphical User Interface (GUI) is designed. GUIs are the core medium for HRI, although sometimes physical devices such as haptics [23] or Brain-Computer Interfaces (BCI) [24] are used. It is necessary for interfaces to effectively convey the environment and information around the robot in order for the user to make an informed choice. Recently, there has been an expansion in modes of interaction, drawing from Augmented Reality (AR) and Virtual Reality (VR). These modes have been investigated as methods of demonstrating skills to robots [25]. In [26], a VR interface is compared with a 2D keyboard and mouse interface. It is found that generally, keyboard and mouse interfaces lead to better task performance, despite the natural motion afforded by VR. Following these results, we design our interface to use keyboard and mouse, while kinesthetic teaching is used for robot demonstrations.

This work combines preference learning and estimation with Learning from Demonstration techniques. We hypothesize that humans have latent preferences for how skills should be performed. However, these preferences may not be fully reflected in the skill demonstrations, which are often suboptimal. Therefore, our goal is to enable robots to estimate and potentially learn human preferences through interaction. To estimate human preferences, our framework (shown in Figure 1) from the demonstration of a task, and then iteratively interacts with the user to find the human's preferred way of reproducing the modeled task. This requires a skill learning representation capable of producing reproductions of an encoded skill in various ways through multi-objectivity. We achieve this through a novel formulation for LfD consisting of an internal optimization problem that discovers optimal reproductions of the skill, integrated with a multi-objective meta-optimization formulation that generates variations of the skill reproduction by trading off the set of objectives. After a few exploratory interactions with the skill learning model, we estimate the user's preference. The discovered human preferences then can be compared across users and tasks. Our contributions include: (a) a novel formulation for LfD according to preference and a multi-objective representation of reproductions in preference space, (b) design of a GUI for human-in-the-loop interactions to find preference, and (c) analysis of preferences with respect to tasks and other users. We have tested our framework with 20 users across 3 different tasks. We have also validated our assumptions used in these human-in-the-loop preference interactions in a

separate study of 10 users, finding that users strongly prefer optimal robot executions. We find that while different users and tasks do not often overlap in preference space, there are certain attributes, such as specific amounts of smoothness, which show strong trends.

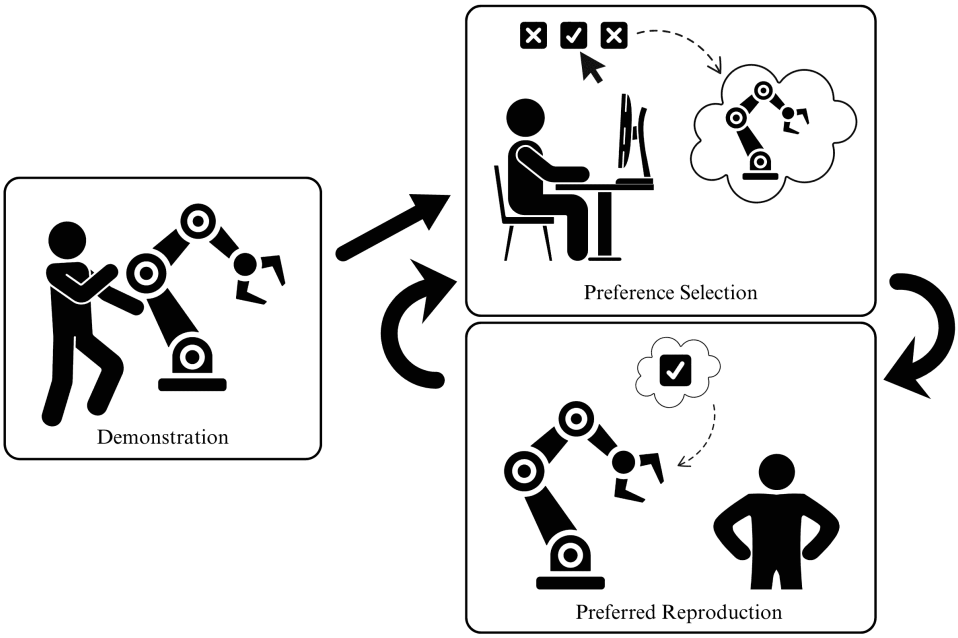

Figure 1. Given a set of demonstrations, our framework builds a skill model and reproduces the skill according to multiple objectives. We then estimate user preference from feedback on various reproductions of the skill over time.

2. Methods

In this section, we formalize a skill learning algorithm that enables robots to learn from human demonstrations. The demonstrations are encoded using an energy-based representation that allows for a variety of reproductions. To allow for flexible reproduction of the skill under varying objectives, we introduce a meta-optimization framework that wraps around the learning algorithm. This framework enables the generation of the modeled skill while selectively emphasizing different objectives. We also discuss the GUI which we design to facilitate interactions with users to capture preference, and how we estimate user preference from interactions with our GUI.

2.1. Skill Learning from Demonstrations

Reproducing tasks aligned with user preferences requires a flexible LfD representation. There are several popular LfD representations that are unable to perform to the specifications of our framework, such as Dynamic Movement Primitives [27,28], which can change their behavior according to several parameters but cannot replicate jerkiness in demonstrations, or Probabilistic Movement Primitives [29], which require multiple demonstrations for best results. One LfD method of interest is Multi-Coordinate Cost Balancing [30], which creates reproductions by balancing between several differential coordinate frames, where the coordinate frames could represent arbitrary features. Unfortunately, this method cannot introduce variable smoothness. One method that can have finely-tuned variable smoothness is elastic maps [31–33].

Building upon this idea, we propose a novel skill learning method that uses a set of (elastic map-like) energy terms to generate reproductions that balance between differential coordinates and motion smoothness, controlled by a set of tuning parameters. Notably, this approach differs from previous methods using elastic maps due to our inclusion of multi-coordinate energy terms and tuning methods. To balance similarity in differential coordinate frames and motion smoothness, we formulate the following energy terms:

$$u_1 = (\mathbf{x}-\mathbf{y})^\top \mathbf{I}^\top \mathbf{I}(\mathbf{x}-\mathbf{y}), \tag{1}$$

$$u_2 = (\mathbf{x}-\mathbf{y})^\top \mathbf{E}^\top \mathbf{E}(\mathbf{x}-\mathbf{y}), \tag{2}$$

$$u_3 = (\mathbf{x}-\mathbf{y})^\top \mathbf{R}^\top \mathbf{R}(\mathbf{x}-\mathbf{y}), \tag{3}$$

$$u_4 = \mathbf{y}^\top \mathbf{E}^\top \mathbf{E}\mathbf{y}, \tag{4}$$

$$u_5 = \mathbf{y}^\top \mathbf{R}^\top \mathbf{R}\mathbf{y}, \tag{5}$$

where u_1 is the Cartesian approximation energy, defined as the difference in Cartesian space between the demonstrations $\mathbf{x} \in \mathbb{R}^{n\times d}$ and the reproduction $\mathbf{y} \in \mathbb{R}^{n\times d}$, where $\mathbf{I} \in \mathbb{R}^{n\times n}$ is the identity matrix. The Tangent approximation energy, u_2, is defined as the difference in Tangent space between the demonstration data and the reproduction, where $\mathbf{E} \in \mathbb{R}^{n-1\times n}$ is the first-order difference matrix [34]. The Laplacian approximation energy, u_3, is defined as the difference in Laplacian space between the demonstration data and the reproduction, where $\mathbf{R} \in \mathbb{R}^{n-2\times n}$ is the second-order difference matrix [34]. Finally, u_4 and u_5 are the stretching and bending energies, respectively. The approximation energies represent weighted squared distance error terms, promoting adherence of reproductions $\mathbf{y}$ to demonstrations $\mathbf{x}$ in each differential coordinate, where $\mathbf{E}^\top\mathbf{E}$, $\mathbf{R}^\top\mathbf{R}$, and for consistency, $\mathbf{I}^\top\mathbf{I}$ are positive semi-definite matrices. Typically, elastic maps rely on the Cartesian approximation energy as well as stretching and bending energies (u_1, u_4, u_5), shown in Figure 2 [32]. These are constructed by attaching a series of nodes (which can be used as a reproduction) to adjacent nodes and surrounding demonstration data via springs. An optimization minimizing the spring energy between the nodes and modeled data (u_1), the spring energy of stretching between adjacent nodes (u_4), and the spring energy of bending two adjacent spring connections (u_5) results in an optimal reproduction. However, these energies were not enough to capture a wide variety of behaviors, such as maintaining the shape of a demonstration while ignoring its Cartesian space properties. Therefore, we included u_2 and u_3, allowing our method to capture a wide variety of behaviors according to a user's preference. Given a set of demonstrations, we find an optimal reproduction of the skill by solving the following minimization problem:

$$\begin{aligned} &\underset{\mathbf{y}}{\text{minimize}}\ f_0(\mathbf{y}) = \sum_{i=1}^{N} \alpha_i u_i \\ &\text{s. t. } f_1(\mathbf{y}) = \mathbf{C}\mathbf{y} = \mathbf{x}, \end{aligned} \tag{6}$$

that is a convex sum of the energy terms (Equations (1)–(5)) weighted by scalar values $\alpha_i \geq 0$ for $i = 1,\ldots,N$, where N is the number of energy terms ($N = 5$ in our case). The optimization problem includes a set of equality linear constraints $f_1(\mathbf{y})$ represented by the matrix $\mathbf{C}$ for constraining a point along the reproduction $\mathbf{y}_j$ to another point in $\mathbf{x}_j$, the demonstration space. These set of equality constraints has been included to ensure the key points of each task are preserved automatically, which results in reducing user load and increasing task performance. The tasks chosen in our experiment are all defined

by events such as actuating the gripper (e.g., in pick & place or pushing) or reaching some spatial extrema (e.g., in pushing or erasing). To ensure important key points were preserved, we apply constraints to: (i) the start and end points of the demonstrations, (ii) the maximum and minimum points in each axis in Cartesian space, and (iii) any point the gripper is opened or closed.

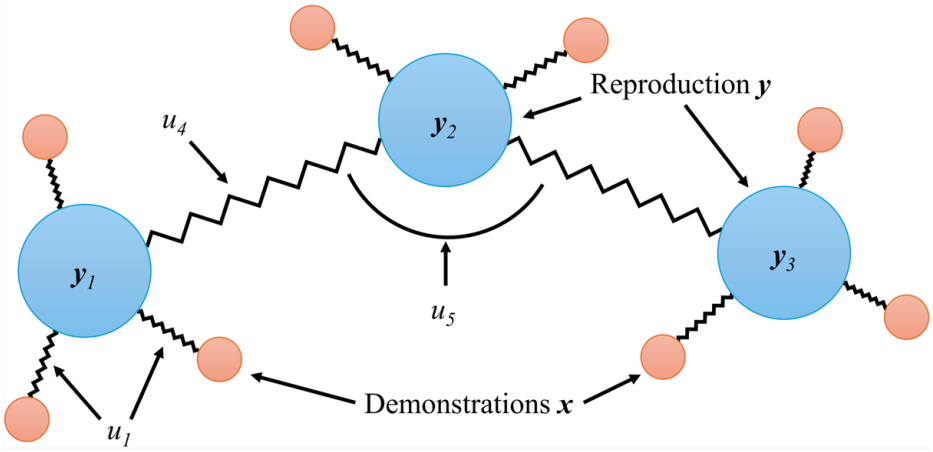

Figure 2. Visualization of energies present in the elastic map. The elastic map finds a reproduction $\mathbf{y}$ which minimizes the spring energies (u_1, u_4, u_5) associated between the reproduction and demonstration(s).

The formulated optimization problem including five quadratic objectives is convex. Therefore, the set of objectives in Equation (6) can be combined into a single quadratic objective:

$$\begin{aligned} U &= (\mathbf{x}-\mathbf{y})^\top \mathbf{Q}(\mathbf{x}-\mathbf{y}) + \mathbf{y}^\top \mathbf{A}\mathbf{y} \\ &= \mathbf{y}^\top (\mathbf{Q}+\mathbf{A})\mathbf{y} - 2\mathbf{x}^\top \mathbf{Q}\mathbf{y}, \end{aligned}$$

where $\mathbf{Q} = \alpha_1 \mathbf{I}^\top \mathbf{I} + \alpha_2 \mathbf{E}^\top \mathbf{E} + \alpha_3 \mathbf{R}^\top \mathbf{R}$ and $\mathbf{A} = \alpha_4 \mathbf{E}^\top \mathbf{E} + \alpha_5 \mathbf{R}^\top \mathbf{R}$ are symmetric positive semi-definite matrices. Note that the constant term $\mathbf{x}^\top \mathbf{Q}\mathbf{x}$ has been removed from the objective because it is independent of $\mathbf{y}$. We obtain the following convex optimization formulation which is a quadratic program:

$$\begin{aligned} &\text{minimize } \mathbf{y}^\top (\mathbf{Q}+\mathbf{A})\mathbf{y} - 2\mathbf{x}^\top \mathbf{Q}\mathbf{y} \\ &\text{s. t. } \mathbf{C}\mathbf{y} = \mathbf{x}, \end{aligned} \tag{7}$$

We use the Lagrange multiplier technique to construct the Karush-Kuhn-Tucker (KKT) system:

$$\begin{bmatrix} 2(\mathbf{Q}+\mathbf{A}) & \mathbf{C}^\top \\ \mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \lambda \end{bmatrix} = \begin{bmatrix} 2\mathbf{Q}\mathbf{x} \\ \mathbf{x} \end{bmatrix},$$

where λ is the Lagrange multiplier. This KKT system can be solved efficiently to find the saddle point $(\mathbf{y}^*, \lambda^*)$ on the Lagrangian.

2.2. Multi-objective meta-optimization

The optimal solution of Equation (7) depends on $\mathbf{A}$ and $\mathbf{Q}$ matrices that are functions of the weight parameters as $\mathbf{A}(\alpha_4, \alpha_5)$ and $\mathbf{Q}(\alpha_1, \alpha_2, \alpha_3)$. By adjusting the α_i parameters, we can reproduce the encoded skill according to various preferences. Based on previous literature and relevant background on human movement and biomechanics [35], we selected three objectives as the main dimensions in

which user preference may be captured in association to robot task execution: (i) spatial similarity, denoted as O_{SSE}, measured by Sum of Squared Error between the reproduction and demonstration, $O_{\text{SSE}} = (\mathbf{x}-\mathbf{y})^\top(\mathbf{x}-\mathbf{y})$, (ii) angular distance, denoted by O_{ANG}, between the reproduction and demonstration, which determines the similarity in shape between the reproduction and demonstration as a cosine distance, $O_{\text{ANG}} = (1/\pi)\sum_{i=1}^{n-1}\cos^{-1}((\mathbf{a}\cdot\mathbf{b})/(\|\mathbf{a}\|_2\|\mathbf{b}\|_2))$, where $\|\cdot\|_2$ is the L^2-norm, $\mathbf{a} = \mathbf{y}_{i+1} - \mathbf{y}_i$ and $\mathbf{b} = \mathbf{x}_{i+1} - \mathbf{x}_i$, or the tangent of $\mathbf{y}_i$ and $\mathbf{x}_i$, respectively, and (iii) smoothness, denoted as O_{SMT}, of the reproduction, which is a measure of the total jerk present in the reproduction, $O_{\text{SMT}} = \int \dddot{\mathbf{y}}\, dt$. These objectives relate to the energies presented in Equations (1)–(5). Sum of Squared Error is directly the Cartesian approximation energy, so optimizing O_{SSE} would lead to a large penalty for u_1, represented by α_1. Angular distance penalizes difference in shape, which is captured by the Laplacian approximation energy (u_3). Smoothness is also represented by the energy terms for stretching and bending energies, and (assuming smooth demonstrations) the Tangent approximation energy (u_2, u_4, and u_5). Therefore these objectives can be changed as a result of the changes in the weighting parameters $\alpha_1, \alpha_2, ..., \alpha_5$. These objectives are often used as metrics for evaluating LfD representations [36,37]. Additionally, LfD methods often aim to find some optimal tradeoff between these objectives [38,39]. In [38], a method is proposed which uses a Laplacian transform to preserve the shape properties of a trajectory while adapting to novel constraints. Alternatively, in [39], the authors explicitly formulate a tradeoff between minimizing jerk (maximizing smoothness) and maximizing spatial accuracy in reproducing the given demonstration. However, these objectives have not been used for evaluation of human preferences. To ensure these objectives can accurately represent preferences, we perform a validation experiment described in Section 2.5. Note that other objectives could be used, or more objectives in addition to the three used here. However, the more objectives, the slower the meta-optimization. Therefore, we limit our objectives to these three as they are still able to provide a wide array of behaviors and ability to capture human preferences while not overcomplicating the meta-optimization.

We use these three objectives to form a multi-objective optimization problem with the decision variables $\alpha_1, \alpha_2, ..., \alpha_5$. Notably, this is a meta-optimization on Equation (7), set up as the following bounded unconstrained problem

$$\begin{aligned} &\text{minimize } \mathbf{g}(\mathbf{z}) = [O_{\text{SSE}}(\mathbf{z}), O_{\text{ANG}}(\mathbf{z}), O_{\text{SMT}}(\mathbf{z})]^\top \\ &\text{s. t. } \mathbf{z}_{\min} \leq \mathbf{z} \leq \mathbf{z}_{\max}, \end{aligned} \tag{8}$$

where $\mathbf{z} = [\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5]$ is the set of decision variables, $\mathbf{g}^* \in \mathbb{R}^{M\times 3}$ is the resulting optimal objective values with corresponding optimal parameters $\mathbf{z}^* \in \mathbb{R}^{M\times N}$, and M is the population size. In our experiments, we found that the decision vector bounds of $z_i \in [0, 10^{10}]$ for $i = 1, \ldots, 5$ work well. We solve the multi-objective optimization using the Non-dominated Sorting Genetic Algorithm II (NSGA-II) [40]. NSGA-II creates a series of populations according to constraints which attempt to optimize the given problem, and at each iteration the best solutions (of current and previous populations) are used to generate the next iteration's population. Additionally, more diverse (but still well-suited) solutions are maintained for future generations. This genetic algorithm was chosen as we wish to find a set of options along the Pareto front to present to the user, and NSGA-II generates a diverse set of Pareto optimal solutions quickly. We also tested using NSGA-III [41,42] which is suited towards many-objective optimization problems,

but found this algorithm provided roughly the same solutions with a longer runtime. The algorithm was set to run for 1000 evaluations with a population size of 100.

Preference space: We define the space of all Pareto optimal solutions of Equation (8) as the preference space. As shown in Figure 3, with three objectives, the preference space can be represented as a unit cube. We assume that the user's preferences which lie within this cube can be found through a series of queries generated by trading off the three objectives, $[O_{\text{SSE}}(\mathbf{z}), O_{\text{ANG}}(\mathbf{z}), O_{\text{SMT}}(\mathbf{z})]$. Note that this preference space could incorporate more objectives and be extended into higher dimensions. For the purposes of this work, we have used 3 objectives, and visualize this preference space as a unit cube.

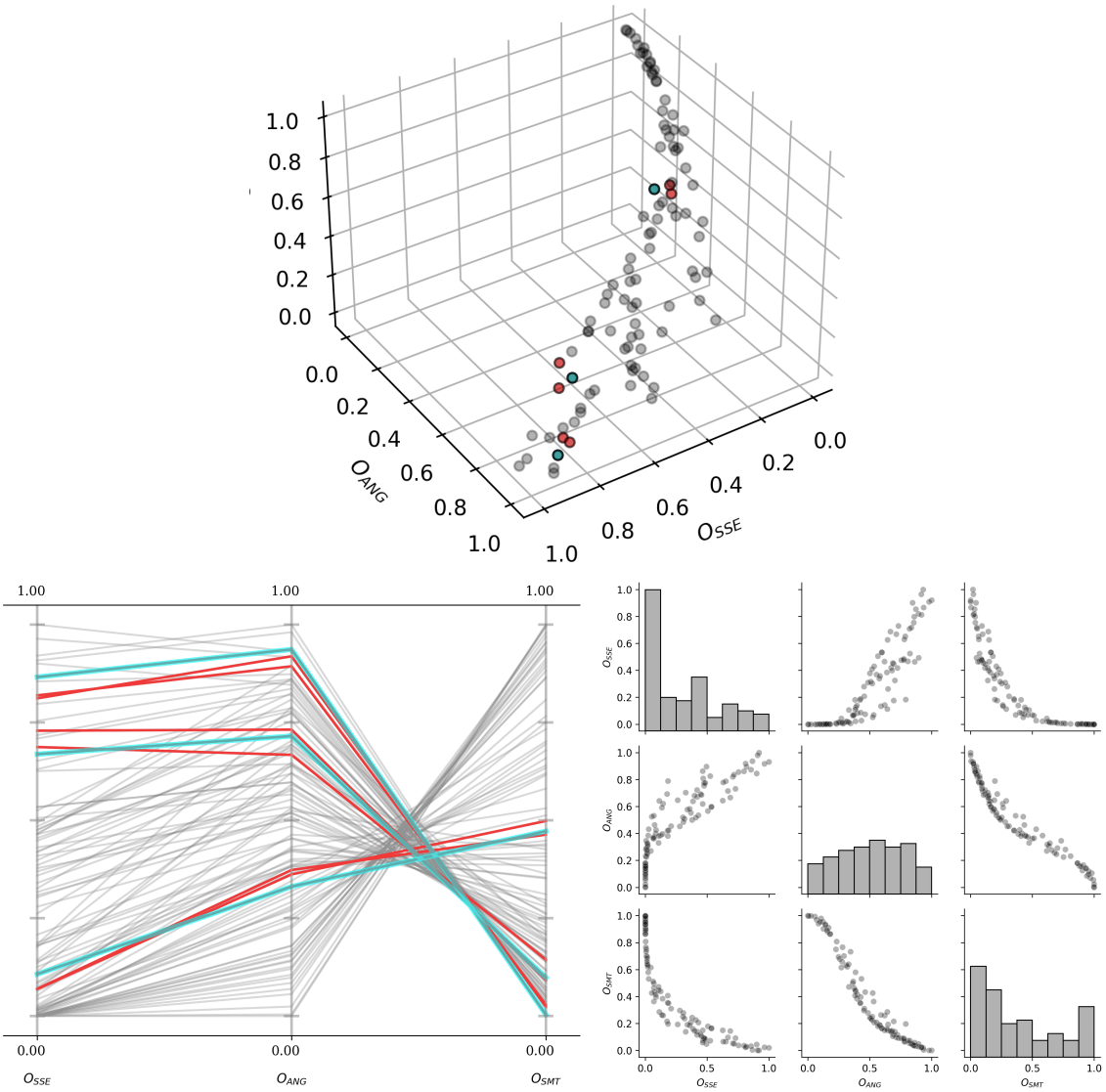

Figure 3. Visualization of options presented to user. Top (3D scatterplot) and bottom left (parallel plot): all possible solutions (Pareto front) are shown in gray, the options shown on the first screen (broad preference) are shown in blue, and the options shown on the second screen (refined preference) are shown in red. Bottom right: 2D projection of the Pareto front options, showing distributions across each objective.

2.3. *User interface design*

To evaluate user preference, we designed a GUI, as seen in Figure 4, which presents several trajectory options to the user prompting them to select their preferred options. Three options are presented to the user to allow for a selection diverse trajectories while not overwhelming the user with too much on the GUI [43]. Each option includes a 3D plot of the reproduction along with the demonstration, a series of performance evaluations (corresponding to objectives used in Equation (8) measured with the current reproduction), and options to select or preview the trajectory. A robot visualization using RViZ is presented at the bottom of the screen. If the user selects the "Preview Option" button for any option, the simulated robot executes the selected trajectory.

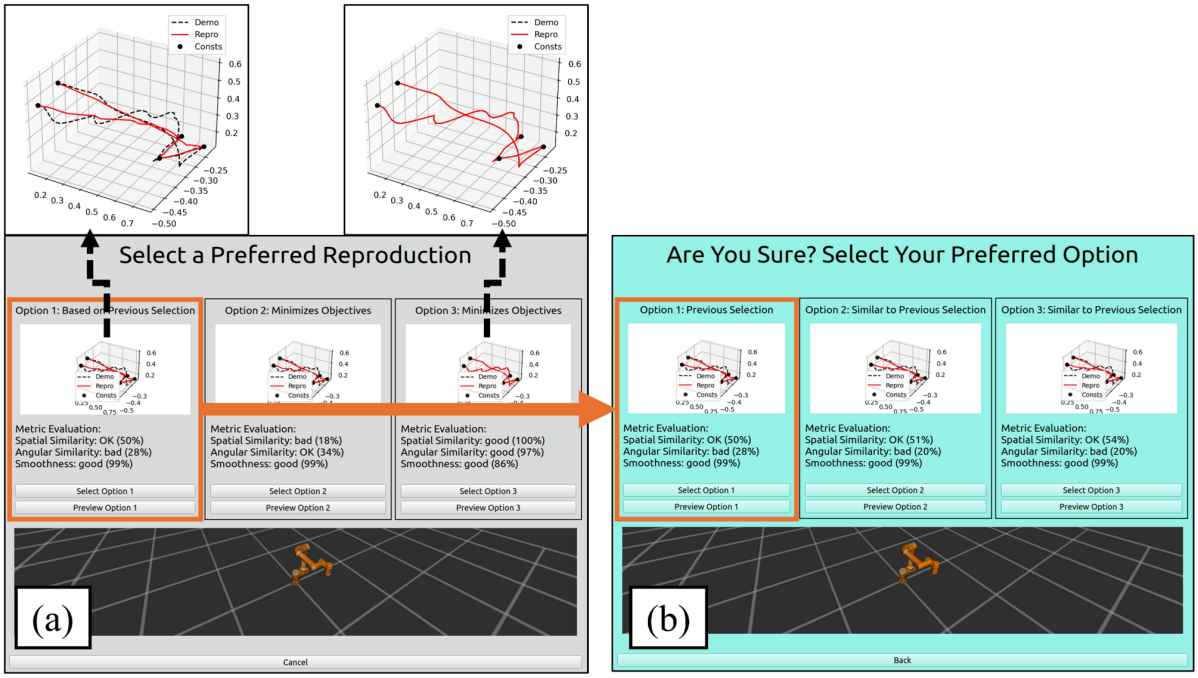

Figure 4. The designed Graphical User Interface (GUI). In **(a)** Users are first shown three different trajectory options in preference space. Upon selecting one of these options (orange box), they are taken to **(b)** where they are presented with three options in preference space that are similar to the selected option.

Once an option is selected on the screen (a), the interface moves to a screen (b) which has the same layout as the first, but presents options similar in preference space to the selected trajectory. A user may choose to go back to the screen (a), or if they select an option, the robot will begin executing the selected movement. Over time, the interface learns the motion preferences of the user, and presents options closer to their conveyed preference.

2.4. *User preference determination*

After finding a set of optimal solutions using Equation (8), three of these solutions are shown to the user when selecting a reproduction. Our design decision behind showing the user 3 options is Miller's Law [44], which states that about 7 options is best for user decisions. We employ a 2-stage selection process, with three options shown at each stage. We employ several algorithms in order to ensure a variety of behaviors are shown to the user for selection, while also allowing for fine-tuning of the user's preference.

An example Pareto front solution can be seen in Figure 3, with a 2D projection of the Pareto front and distributions across each objective. Internally, every reproduction (*i.e.*, a Pareto efficient solution) is normalized to be measured on a scale of $[0,1]$ for each of these objectives (lower is better for all objectives) to create the preference space. To estimate the user's preference, we use an Exponential Moving Average (EMA) with bias correction [45], as

$$p_i = \frac{\beta p_{i-1} + (1-\beta)s_i}{1-\beta^i}, \tag{9}$$

where for the i-th iteration, p_i is the estimated preference, s_i is the selected preference, and β is a tuning parameter. This results in the newly estimated preference p_i as a shift from the previously estimated preference p_{i-1} towards the newly selected preference s_i by some amount according to β. This moving average is used as user preference may shift between tasks, and EMA can quickly respond to shifts.

In order to accurately assess preference, the user must be shown a variety of options to select from. To determine which options were shown in each selection screen, we employ two different algorithms. The first algorithm, used for screen (a), aims to show a variety of options to get a general idea of the user's preference, and the second algorithm, used for screen (b), refines the preference selection. We begin by selecting an option to present to the user which most closely matches their current preference, as

$$r_{i,1} = \underset{r_{i,1} \in \mathbf{g}^*}{\arg\min} \|\mathbf{g}^* - p_i\|_2, \tag{10}$$

where $r_{i,1}$ is the first option shown to the user in the i-th iteration and $\mathbf{g}^*$ is all possible selections given by Equation (8). This only works if there is already a known preference, which isn't available on the first iteration. If preference is unknown, we randomly sample the Pareto front, finding probabilities:

$$\pi_{i,1} = \sigma(\|\mathbf{g}^*\|_2) = \frac{e^{\|\mathbf{g}^*_j\|_2}}{\sum_{k=1}^{K} e^{\|\mathbf{g}^*_k\|_2}} \text{ for } j = 1,\ldots,M, \tag{11}$$

where $\pi_{i,1}$ are the probabilities for iteration i that a potential selection $\hat{s} \in \mathbf{g}^*$ will be shown to the user and $\sigma(\cdot)$ is the softmax function. This probability is sampled to find the first option shown on screen as $r_{i,1} \sim \pi_{i,1}$. We then find the second and third options (whether preference is known or not) using:

$$\pi_{i,j} = \sigma\left(\sum_{k=1}^{j-1} \|\mathbf{g}^* - r_{i,k}\|_2\right), \tag{12}$$

where $r_{i,2} \sim \pi_{i,2}$ is sampled before calculating probabilities $\pi_{i,3}$. This probabilistic method finds options to present to the user which are likely different in preference space. Examples of options presented to the user on screen (a) are shown in cyan in Figure 3, with possible trajectories shown at the top of Figure 4. On screen (b), we wish to present options similar to the option selected on the first screen. Therefore, we show the two closest options in preference space to the selected trajectory, shown in red in Figure 3.

We assume preferences lie on the Pareto front and validate this assumption in a separate experiment. We also give the users choices between several objectives, as we believe only optimizing one of the objectives in Equation (8) leads to poor results, as can be seen in Figures 5 and 6. In each figure, we minimize one of each of the three objectives, which leads to reproductions which do not generalize

(for $O_{\text{SSE}}, O_{\text{ANG}}$) or over-generalize (for O_{SMT}). It is only through a combination of objectives (as seen by the Pareto front option, which minimizes all three) that a preferred reproduction of the task can be found. These reproductions are shown in simulation in Figure 5, and the real-world executions are shown in Figure 6. As can be seen from these results, the execution which provides a mix of optimizing the objectives provides the best erasing of the drawn shape. The forces felt on the robot during execution were also recorded, with the maximum and average results shown in Table 1. The executions which minimize O_{SSE} and O_{ANG} record the lowest maximum forces of the executions, but the Pareto front mix of objectives results in the lowest average forces. These results show that using all objectives in Equation (8) provides the best results rather than optimizing for only a single objective.

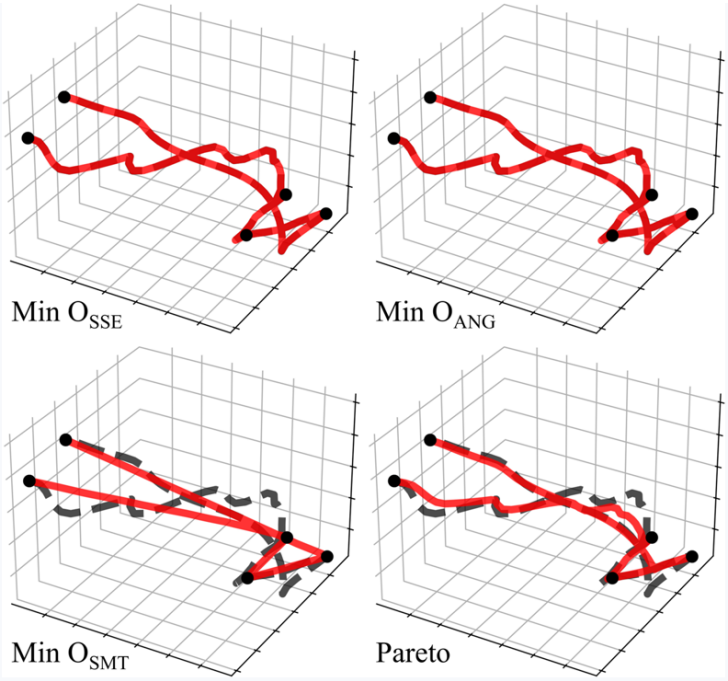

Figure 5. A comparison between the extreme choices for each objective ($O_{\text{SSE}}, O_{\text{ANG}}, O_{\text{SMT}}$) and a Pareto front solution which optimizes all three objectives for a selected erasing demonstration. Real-world executions of these trajectories can be seen in Figure 6.

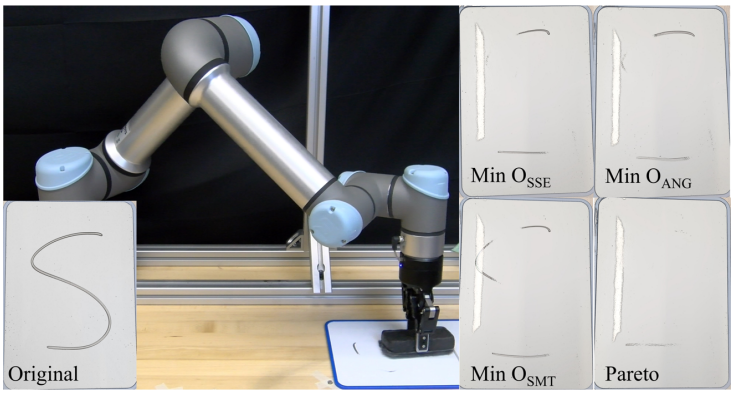

Figure 6. The trajectories shown in Figure 5 executed on the robot. The task was to erase an "S" shape (shown on left). As can be seen of the results on the right, the Pareto front trajectory which minimizes all objectives performs better than minimizing a single objective.

Table 1. Forces felt by the robot across executions which optimize each of the objectives in Equation (8) as well as a Pareto front option which optimizes all objectives.

	Min O_{SSE}	Min O_{ANG}	Min O_{SMT}	Pareto
Max Force (N)	30.49	**30.41**	33.46	32.32
Average Force (N)	4.60	4.44	4.47	**3.77**

2.5. Validating user preferences

In our method described above, we make several key assumptions which are necessary to be validated. The first is our assumptions that users prefer reproductions which lie on the Pareto front as opposed to Non-Pareto solutions. The second is our selection of objectives can accurately represent user preference: spatial similarity, angular similarity, and smoothness. To validate these assumptions, we perform a separate user study. In this experiment, we present a user with a pre-recorded demonstration and 5 options for reproductions: one which maximizes spatial similarity, one which maximizes angular similarity, one which maximizes smoothness, one which is an optimal mix of the three objectives in Equation (8), and one sub-optimal solution which does not lie on the Pareto front. Users are asked to rank these solutions from most to least preferred using the GUI shown in Figure 7. The order in which these reproductions are shown to users is randomized to avoid biases such as the ballot order effect. Users are allowed to preview the reproductions through 3D plots, RViZ visualizations, and executions on the real robot as many times as they wish. Once the user is satisfied with their rankings, their ordering is recorded and the experiment concludes.

Figure 7. The GUI used to allow users to rank their preferred reproductions. Users were asked to rank five reproductions, three of which minimized each objective in Equation (8), one which mixes these objectives similar to user preferences, and one sub-optimal reproduction which does not lie on the Pareto front.

3. Results

3.1. Experimental setup

All experiments were performed in a lab with a Universal Robots UR5e 6DOF manipulator arm equipped with a Robotiq 2f85 gripper (as seen in Figure 8). First, users were brought into the lab and asked to perform one of the following tasks:

- Pick & place—picking a cup from a dishwasher and placing it next to the robot,

- Pushing—closing a dishwasher drawer, or
- Erasing—erasing a shape drawn on a whiteboard next to the robot.

Task order and distribution were counterbalanced between participants. The robot started in a default position for every demonstration and execution, but users could end their demonstrations at an arbitrary position. All demonstrations were captured through kinesthetic teaching. Users were allowed as many attempts as they wished to perform a demonstration that they deemed satisfactory. After each demonstration or execution of the task, the workspace was reset. Both experiments were approved by an IRB, and all participants were compensated for their time.

(a) Pick & Place (b) Pushing (c) Erasing

Figure 8. The three tasks used in our user study: **(a)** pick & place; **(b)** pushing; and **(c)** erasing letter "S".

3.2. User study on preference ranking

We validate our assumptions as described in Section 2.5. In this experiment, users are not asked to give a demonstration of one of the three tasks in Figure 8 and it is instead provided for them. The GUI (shown in Figure 7) was used to capture rankings between reproductions which capture different preferences. This experiment consisted of 10 participants. Most importantly, this experiment seeks to verify that users prefer optimal (Pareto) solutions to suboptimal ones. The rankings in Figure 9 show box-and-whisker plots of the rankings, and it can be seen that the Non-Pareto (NP) solution is significantly the least-preferred option ($p < 0.005$). In fact 8 out of 10 users ranked it as such. Notably, the only

other reproduction to be ranked as least-preferred was the reproduction which maximized smoothness. Both users who ranked the smooth reproduction lower were given an Erasing task, where spatial and angular similarity are more important. This validates our assumption that users strongly prefer optimal reproductions, and therefore we ignore all Non-Pareto solution in our study for exploring user preference.

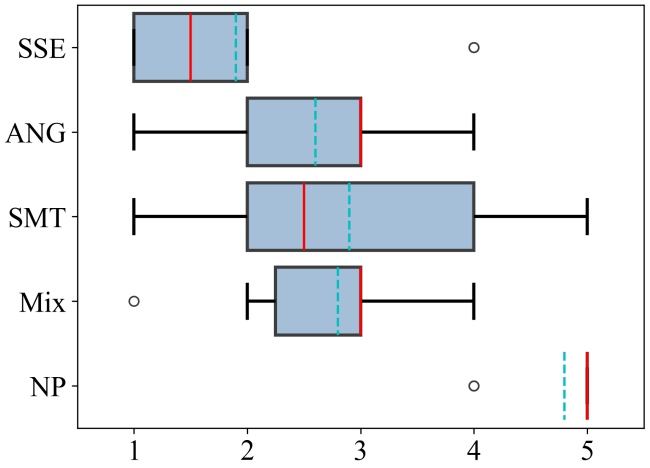

Figure 9. Rankings for validating preference of users. Users were asked to rank five reproductions of a task: three of which respectively minimized each of the objectives in Equation (8), one which provides a balanced mix of optimizing these objectives (Mix), and one suboptimal reproduction which does not lie on the Pareto front (NP). For all box-and-whisker plots, the median is shown with a solid red line, and the mean is shown with a dashed blue line. Whiskers are 1.5 interquartile range.

Additionally, we used this experiment to verify that our selection of objectives (spatial similarity, angular similarity, and smoothness) can accurately represent user preferences. If any of these objectives could not capture user preferences, we would expect it to be ranked lower than the other objectives, which is not seen. In fact, reproductions which minimize each objective individually as well as a mix of the objectives are each ranked by different users as their most-preferred. However, it could be the case that none of the objectives we selected are reflecting user preferences, and optimal solutions are simply better than non-optimal ones. To ensure the presented reproductions reflected user preferences, we conducted an exit survey with the experiment. One of the questions of this survey was "Do you think any of the trajectories accurately represented your preference for the task? (1: Not accurate, 5: Very Accurate)" Participants indicated the trajectories were "mostly accurate" representations of their preference ($\bar{\mu} = 3.90, \sigma = 0.87$). In fact, all responses indicated trajectories were "accurate" or better at reflecting their preference. Therefore, our selected objectives can be used for representing and finding user preferences.

3.3. User study on preference exploration

Next, we allow users to explore their individual preferences for the tasks shown in Figure 8. The user was shown three options from our GUI (shown in Figure 4), and asked to select their preferred option. The user can preview each option by looking at the resulting trajectory the robot would take through the workspace or with a visualization of the robot motion using planning with MoveIt! and RViZ. Once the user selects one of these options, they are shown screen (b) with the selected option and two other

similar options. Once again, they are able to investigate and preview each option. Upon selecting an option in screen (b), the robot begins execution, and the user is asked to watch the robot while it executes. Movements are executed through planning on MoveIt! and through robot's low-level controller. Once the execution completes, the robot returns to the beginning of the demonstration, and the interface resets to screen (a), but updates the user preference estimation. After a series of 5 reproductions, another of the three tasks was demonstrated and a series of 5 reproductions found. The interface used the preference learned from the first task to estimate the user's preference for the second task.

3.4. Analysis

We conducted an IRB-approved experiment with 20 participants (10 female, 10 male, ages $\bar{\mu} = 23.3, \sigma = 5.39$) using the experimental setup described above and shown in Figure 8. The goal of this experiment was to discover user preferences and analyze similarities and differences in preference across users and tasks.

Task-specific preferences: First, we examine if task preference is shared across humans. If preferences for tasks were shared, this would mean that for a specific task, there is a general way humans prefer robots to perform the task. We plot all user preference selections throughout the experiment in Figure 10. Shared task preference would be indicated by a tight grouping of same-colored points in Figure 10. As can be seen in the figure, no such grouping exists. This indicates that different humans do not share the same preferences for tasks. Measuring the maximum distances between preferences for the same task gives 1.33, 1.39, and 1.31 for pick & place, pushing, and reaching, respectively (maximum is 1.73). This indicates a similar spread across the preference space for each of the three tasks.

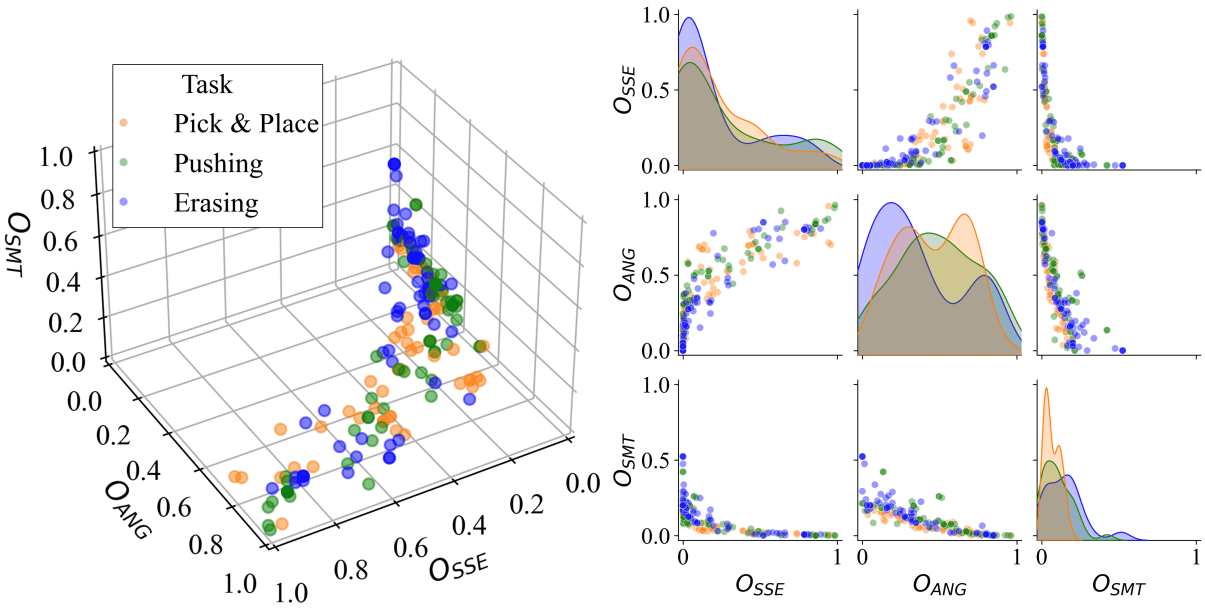

Figure 10. User preference selections according to task, with distributions for each task shown across each objective. Preferences selected during the pick & place task are shown in orange, pushing are shown in green, and erasing are shown in blue. There is no discernible difference of preference between tasks.

Additionally, we can look at the means (shown in Table 2) and covariances (shown in Table 3) for the selected preferences. If covariance of one of the features used in preference determination is low, that means there is a strong preference towards the selected value of that specific feature. The lowest

covariance value is the variance of smoothness in the pick & place task (shown in orange in Figure 10), with a variance of 0.003 and mean of 0.07. This indicates that users preferred a high amount of smoothness in the reproductions of the pick & place task. The variance of smoothness in the pushing task (shown in green in Figure 10) is also low, with a variance of 0.009 and mean of 0.10. This also indicates a preference for higher smoothness, although the preference is not as strong in the pushing task and the amount of smoothness preferred is not as much. A very high covariance value conversely indicates that feature was not as important in that task. The highest variance seen was the variance of spatial similarity in the pushing task, with a variance of 0.111 and mean of 0.27. This indicates that users did not strongly trend towards a specific spatial similarity for this task. We can also compare the means of the same objective across different tasks, using t-tests to determine if there are statistically significant differences between means. If there are, it would indicate that humans do prefer different objectives across different tasks, but just for that objective. Starting with spatial similarity, we do not find support for significant difference in means according to the t-test. In angular similarity, there is a significant difference ($p < 0.05$) in the means between erasing and the two other tasks, but no support for a significant difference between pick & place compared to pushing. This indicates that angular similarity is significantly more preferred in the erasing task, likely in order to properly erase the shape drawn (most participants guided the robot to erase the shape in an "S" pattern similar to how the shape was drawn). Finally, for smoothness, we find significant differences ($p < 0.05$) between all three tasks. This indicates a different level of smoothness is preferred for each task, with the most smoothness preferred in the pick & place task and least smoothness preferred in the erasing task. While humans do not exhibit a specific preference across all three tasks for the features used, there are preferences for specific objectives across the three tasks. Most notably, smoothness varies for each task, meaning humans exhibit strong preferences towards specific levels of smoothness.

User-specific preferences: Next, we examine whether the same user shares preference in different tasks. We look at how users' preference selections change between tasks, illustrated for 4 selected users in Figure 11 where each task is shown with a different color. Preferences are shown in 2D using a projection with Principal Component Analysis (PCA), with a background coloring found according to a Support Vector Classifier, and the 3D representation in objective space shown in the corner. Some users seemingly have similar preferences between tasks (users (a) and (b)) while others have very different preferences (users (c) and (d)). For each user, we calculated the minimum, maximum, and average distance between the preference selections of the first and second task. The results of these calculations are shown on the right in Figure 11, along with markings corresponding to the four users shown. Despite some similarities, each user varies widely from each other. User (a) has lower mean and average distances, but a high maximum distance, indicating significant overlap but inconsistent preference, as seen by the outlier in Figure 11a. User (b) has two tight but separated groups, with the lowest max and average distances among any user (but a higher minimum than even user (a)). User (c) has the least overlap of any user with the highest minimum distance between task preferences, making for distinct preferences between tasks. User (d) has the highest average distance between tasks, but also has some overlapping preference. As shown in the figure, across all users, the minimum distances are quite small, while the maximum preferences are quite large, indicating some overlap between preference selections, but with a large spread. The average of the average distances is 0.36 (the maximum possible distance is 1.73), which does not indicate preference

transfers between tasks for the same user. We also look at t-tests for finding statistically significant differences in the means between a user's selections in the first task compared to the second task. We find that only 7 users had significantly different means for the spatial similarity objective, 6 for the angular similarity objective, and 7 for the smoothness objective. Overall, 4 users showed differences across all three objectives, while 10 users had no support for a significant difference across any of the objectives. This shows that different users vary widely on whether preference is shared between tasks, highlighting the need for individual preferences and allowing users the ability to customize robot reproductions to their preference.

Table 2. Overall and task-specific means for each objective (lower indicates stronger preference).

	O_{SSE}	O_{ANG}	O_{SMT}
Pick & Place	0.24	0.46	0.07
Pushing	0.27	0.48	0.10
Erasing	0.21	0.36	0.15
Overall	0.24	0.43	0.11

Table 3. Overall and task-specific covariances for each objective (lower indicates a more consistent preference).

		O_{SSE}	O_{ANG}	O_{SMT}
Pick & Place	O_{SSE}	0.076	0.049	−0.011
	O_{ANG}	0.049	0.052	−0.011
	O_{SMT}	−0.011	−0.011	0.003
Pushing	O_{SSE}	0.111	0.077	−0.023
	O_{ANG}	0.077	0.070	−0.019
	O_{SMT}	−0.023	−0.019	0.009
Erasing	O_{SSE}	0.078	0.072	−0.025
	O_{ANG}	0.072	0.081	−0.029
	O_{SMT}	−0.025	−0.029	0.016
Overall	-	0.088	0.071	0.010

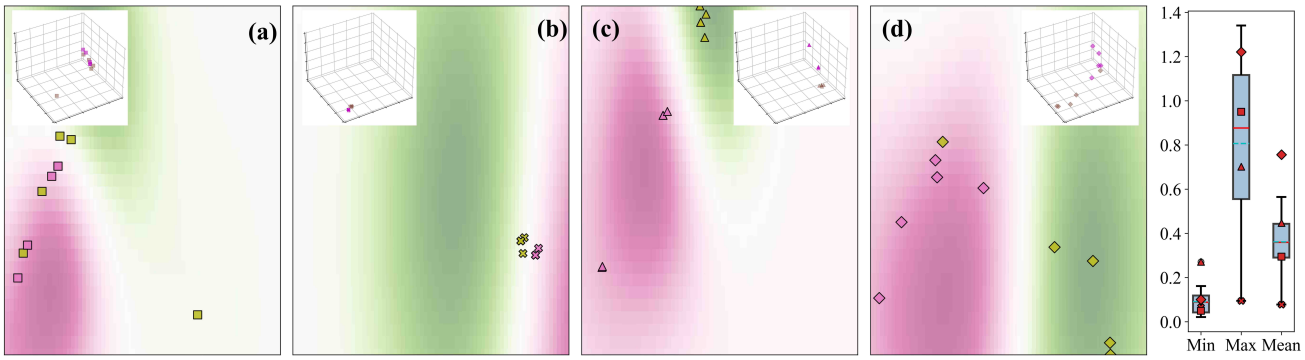

Figure 11. User preference between tasks for selected users, shown in 3D preference space as well as projected onto 2D space using PCA. Preference selections in the first task are shown in green, while preference selections for the second task are shown in magenta. Each user is represented with a different marker. Some users **(a, b)** have similar preference between tasks, while others **(c, d)** show distinct preferences. Distances between clusters of preference selections for the first and second tasks are shown on the right, with markers for the specific users shown in **(a)–(d)**.

User-demonstrated preferences: Finally, we investigate if users demonstrate tasks according to their preference. If users prefer their demonstrations, they should select options with high spatial similarity. We measure overall means and covariances across all users and tasks, which are shown in Tables 2 and 3, respectively. If users demonstrate tasks according to their own preferences, there should be a high preference for spatial similarity and a low variance for that objective. The mean spatial similarity is 0.24 with a variance of 0.088. This indicates some preference for spatial similarity, but this preference is not always present nor consistent. In fact, spatial similarity is the least consistent objective across preferences as it has the highest variance. Alternatively, we find that smoothness has a very high preference and low covariance with a mean of 0.11 and covariance of 0.010. This indicates that users do not perform demonstrations according to their preferences, and reproductions should be adjusted accordingly, mainly finding reproductions which are smoother than demonstrations.

Accurate user preferences: Once again, to ensure the presented reproductions reflected user preferences, we conducted an exit survey with the experiment. Here, we asked participants two questions of interest, measuring if suggestions from the interface were useful and if they felt their preference had been accurately found. For usefulness, we asked "Were the suggestions from the interface useful? (1: Not Useful, 5: Very Useful)" Participants indicated the suggestions were "more than useful" ($\bar{\mu} = 3.70, \sigma = 0.92$), with only 2 participants indicating they felt the suggestions were "less than useful." This shows that participants were able to use the interface's suggestions to explore alternative preferences. As for preference accuracy, we asked "Do you think the interface accurately found your preference for the task? (1: Not accurate, 5: Very Accurate)" Generally, participants felt that the found preference was either "accurate" or "mostly accurate" ($\bar{\mu} = 3.50, \sigma = 1.00$), although there was higher variation in responses than in our preference ranking experiment. This variation likely comes from the preference adjustment between tasks, which our method was not able to capture well for all participants.

4. Conclusions

In this work, we evaluated human preferences of robot reproductions using a human-in-the-loop framework, where the human first demonstrated a task, then worked with the robot through an interface to find their preferred reproduction. After some time, the task was switched, but their predicted preference carried over. The user's new preference was then re-estimated for the new task. We measured if human preference transfers between tasks, across users, and if user demonstrations follow user preference.

There are a few key takeaways from our experiments. The first is that users strongly prefer optimal reproductions to suboptimal ones. Since user demonstrations are often suboptimal themselves [8], this means optimal motion planning methods are necessary for user-preferred robotic movements. Next, we find that users and tasks are highly individualistic, meaning that user preferences vary highly from user to user and task to task in a general sense. Therefore, robot motions should be tailored to each user for each task, highlighting the need for preference estimation and selection frameworks such as the one presented here. We do find some trends among preferences across tasks and users. Smoothness was highly preferred and consistently preferred across all tasks. We also find that each task has a different amount of preferred smoothness, meaning that preferences for tasks across users cannot be transferred. For Learning from

Demonstration methods, this could mean that a single hyperparameter for tuning smoothness is all that is needed to adjust to user preferences.

There are several opportunities for future work. One limitation of this work is that preferences may not have adjusted fast enough when performing a new task (our initial assumption was that preference would mostly carry over, and large changes in preference were unlikely). Therefore, the ability to switch preference rapidly could be improved, which could also be a viable option for preference switching on-the-fly for different tasks. Also, our method for finding preference is not limited to the LfD method used here, nor the objectives of our meta-optimization. Further studies could be done using alternative LfD methods, or alternative objectives. As a framework for finding user preferences, our method works with any LfD method which can adjust behavior according to parameters and with any objectives which can be formulated into a preference space. In this work, we were somewhat limited in our task selection according to the LfD representation method used. Should the LfD method used have the ability to handle longer or more complicated tasks (such as force-inclusive tasks), preferences could be explored in these novel tasks as well. Alternatively, giving users more control in their options, such as using continuous sliders to adjust preference instead of presenting discrete options, could lead to more accurate user preferences.

Supplementary data

An implementation for the GUIs used in our experiments can be found at: https://github.com/brenhertel/Preference-GUIs.

Data availability statement

The data that support the findings of this study are not publicly available due to privacy, but are available from the corresponding author upon reasonable request.

Declaration of generative AI and AI-assisted technologies

The authors did not use generative AI or AI-assisted technologies in the writing of this manuscript.

Acknowledgments

This research is supported in part by the National Science Foundation (FRR-2237463).

Authors' contribution

BH: conceptualization, formal analysis, investigation, methodology, software, visualization, writing—original draft preparation, writing—review and editing; TN: formal analysis, investigation, writing—original draft preparation, writing—review and editing; MC: conceptualization, funding acquisition, resources, supervision, writing—original draft preparation, writing—review and editing; RA: conceptualization, funding acquisition, resources, supervision, writing—original draft preparation, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

The authors disclose no knowledge of any potential conflicts of interest.

Ethical statement

The study was performed in accordance with the Declaration of Helsinki and approved by the University of Massachusetts Lowell Institutional Review Board number 24-168, approved November 7, 2024.

References

[1] Von Zitzewitz J, Boesch PM, Wolf P, Riener R. Quantifying the human likeness of a humanoid robot. *Int. J. Soc. Robot.* 2013, 5(2):263–276.

[2] García N, Rosell J, Suárez R. Motion planning by demonstration with human-likeness evaluation for dual-arm robots. *IEEE Trans. Syst. Man Cybern. Syst.* 2017, 49(11):2298–2307.

[3] Torricelli D, Mizanoor RSM, Lippi V, Weckx M, Mathijssen G, *et al.* Benchmarking human likeness of bipedal robot locomotion: state of the art and future trends. In *Metrics of Sensory Motor Coordination and Integration in Robots and Animals: How to Measure the Success of Bioinspired Solutions with Respect to their Natural Models, and Against More 'Artificial' Solutions?*, 1st ed. Cham: Springer International Publishing, 2020. pp. 147–166.

[4] Fink J. Anthropomorphism and human likeness in the design of robots and human-robot interaction. In *Proceedings of International Conference on Social Robotics*, Chendu, China, October 29–31, 2012, pp. 199–208.

[5] Sadigh D, Dragan A, Sastry S, Seshia S. Active preference-based learning of reward functions. In *Proceedings of Robotics: Science and Systems XIII (RSS 2017)*, Cambridge, USA, July 12–16, 2017.

[6] Dennler N, Nikolaidis S, Matarić M. Contrastive learning from exploratory actions: leveraging natural interactions for preference elicitation. In *Proceedings of 2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Melbourne, Australia, March 4–6, 2025, pp. 778–788.

[7] Ravichandar H, Polydoros AS, Chernova S, Billard A. Recent advances in robot learning from demonstration. *Annu. Rev. Control Robot. Auton. Syst.* 2020, 3(1):297–330.

[8] Rana MA, Chen D, Williams J, Chu V, Ahmadzadeh SR, *et al.* Benchmark for skill learning from demonstration: Impact of user experience, task complexity, and start configuration on performance. In *Proceedings of 2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, May 31–August 31, 2020, pp. 7561–7567.

[9] Shah R, Gundotra N, Abbeel P, Dragan A. On the feasibility of learning, rather than assuming, human biases for reward inference. In *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, USA, June 9–15, 2019, pp. 5670–5679.

[10] Basu C, Yang Q, Hungerman D, Singhal M, Dragan AD. Do you want your autonomous car to drive like you? In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, Vienna, Austria, March 6–9, 2017, pp. 417–425.

[11] Dragan AD, Lee KC, Srinivasa SS. Legibility and predictability of robot motion. In *Proceedings*

of 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Tokyo, Japan, March 4–6, 2013, pp. 301–308.

[12] Lemasurier G, Bejerano G, Albanese V, Parrillo J, Yanco HA, *et al.* Methods for expressing robot intent for human–robot collaboration in shared workspaces. *ACM Trans. Hum.-Robot Interact.* 2021, 10(4):1–27.

[13] Rossi S, Coppola A, Gaita M, Rossi A. Human–Robot Interaction Video Sequencing Task (HRIVST) for robot's behavior legibility. *IEEE Trans. Hum. Mach. Syst.* 2023, 53(6):975–984.

[14] Wang A, Fitzgerald T. Legibility-aware learning from corrections. In *Proceedings of 2024 19th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Melbourne, Australia, March 11–14, 2024.

[15] Cakmak M, Thomaz AL. Designing robot learners that ask good questions. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, Boston, USA, March 5–8, 2012, pp. 17–24.

[16] Brown DS, Goo W, Niekum S. Better-than-demonstrator imitation learning via automatically-ranked demonstrations. In *Proceedings of the Conference on Robot Learning*, online, November 16–18, 2020, pp. 330–359.

[17] Jonnavittula A, Losey DP. I know what you meant: learning human objectives by (under) estimating their choice set. In *Proceedings of 2021 IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China, May 30–June 5, 2021, pp. 2747–2753.

[18] Liu W, Zhong J, Wu R, Fylstra BL, Si J, *et al.* Inferring human-robot performance objectives during locomotion using inverse reinforcement learning and inverse optimal control. *IEEE Robot. Autom. Lett.* 2022, 7(2):2549–2556.

[19] Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. *J. Big Data* 2016, 3(1):9.

[20] Josifovski J, Malmir M, Klarmann N, Žagar BL, Navarro-Guerrero N, *et al.* Analysis of randomization effects on sim2real transfer in reinforcement learning for robotic manipulation tasks. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Xi'an, China, October 23–27, 2022, pp. 10193–10200.

[21] Truong J, Chernova S, Batra D. Bi-directional domain adaptation for sim2real transfer of embodied navigation agents. *IEEE Robot. Autom. Lett.* 2021, 6(2):2634–2641.

[22] Zhu Y, Tang Z, Liu Y, Zhuang F, Xie R, *et al.* Personalized transfer of user preferences for cross-domain recommendation. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, Tempe, USA, February 21–25, 2022, pp. 1507–1515.

[23] Thakur S, Armas ND, Adegite J, Pandey R, Mead J, *et al.* A tetherless soft robotic wearable haptic human machine interface for robot teleoperation. In *Proceedings of 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Abu Dhabi, United Arab Emirates, October 14–18, 2024, pp. 12226–12233.

[24] Yang C, Wu H, Li Z, He W, Wang N, *et al.* Mind control of a robotic arm with visual fusion technology. *IEEE Trans. Ind. Inform.* 2017, 14(9):3822–3830.

[25] Boguslavskii N, Zhong Z, Genua LM, Li Z. A shared autonomous nursing robot assistant with dynamic workspace for versatile mobile manipulation. In *Proceedings of 2023 IEEE/RSJ International Conference*

on Intelligent Robots and Systems (IROS), Detroit, USA, October 1–5, 2023, pp. 7040–7045.

[26] LeMasurier G, Tukpah J, Wonsick M, Allspaw J, Hertel B, *et al.* Comparing a 2D keyboard and mouse interface to virtual reality for human-in-the-loop robot planning for mobile manipulation. In *Proceedings of 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*, Pasadena, USA, August 26–30, 2024, pp. 2197–2203.

[27] Pastor P, Hoffmann H, Asfour T, Schaal S. Learning and generalization of motor skills by learning from demonstration. In *Proceedings of 2009 IEEE International Conference on Robotics and Automation*, Kobe, Japan, May 12–17, 2009, pp. 763–768.

[28] Ijspeert AJ, Nakanishi J, Hoffmann H, Pastor P, Schaal S. Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Comput.* 2013, 25(2):328–373.

[29] Paraschos A, Daniel C, Peters JR, Neumann G. Probabilistic movement primitives. In *Proceedings of Advances in Neural Information Processing Systems 26 (NIPS 2013)*, Lake Tahoe, USA, December 5–10, 2013.

[30] Ravichandar H, Ahmadzadeh SR, Rana MA, Chernova S. Skill acquisition via automated multi-coordinate cost balancing. In *Proceedings of 2019 International Conference on Robotics and Automation (ICRA)*, Montreal, Canada, May 20–24, 2019, pp. 7776–7782.

[31] Gorban AN, Zinovyev AY. Principal graphs and manifolds. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, 1st ed. Hershey: IGI Global Scientific Publishing, 2010. pp. 28–59.

[32] Hertel B, Pelland M, Ahmadzadeh SR. Robot learning from demonstration using elastic maps. In *Proceedings of 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Kyoto, Japan, October 23–27, 2022, pp. 7407–7413.

[33] Hertel B, Ahmadzadeh SR. Robot learning using multi-coordinate elastic maps. In *Proceedings of 2025 22nd International Conference on Ubiquitous Robots (UR)*, Texas, USA, June 30–July 2, 2025, pp. 301–306.

[34] Hertel B, Ahmadzadeh SR. Confidence-based skill reproduction through perturbation analysis. In *Proceedings of 2023 20th International Conference on Ubiquitous Robots (UR)*, Honolulu, USA, June 25–28, 2023, pp. 165–170.

[35] Flash T, Hogan N. The coordination of arm movements: an experimentally confirmed mathematical model. *J. Neurosci.* 1985, 5(7):1688–1703.

[36] Huang Y, Rozo L, Silvério J, Caldwell DG. Kernelized movement primitives. *Int. J. Robotics Res.* 2019, 38(7):833–852.

[37] Hertel B, Ahmadzadeh SR. Similarity-aware skill reproduction based on multi-representational learning from demonstration. In *Proceedings of 2021 20th International Conference on Advanced Robotics (ICAR)*, Ljubljana, Slovenia, December 6–10, 2021, pp. 652–657.

[38] Nierhoff T, Hirche S, Nakamura Y. Spatial adaption of robot trajectories based on laplacian trajectory editing. *Auton. Robots* 2016, 40(1):159–173.

[39] Meirovitch Y, Bennequin D, Flash T. Geometrical invariance and smoothness maximization for task-space movement generation. *IEEE Trans. Robot.* 2016, 32(4):837–853.

[40] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm:

NSGA-II. *IEEE Trans. Evol. Comput.* 2002, 6(2):182–197.

[41] Deb K, Jain H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. *IEEE Trans. Evol. Comput.* 2013, 18(4):577–601.

[42] Jain H, Deb K. An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: Handling constraints and extending to an adaptive approach. *IEEE Trans. Evol. Comput.* 2013, 18(4):602–622.

[43] Szafir D, Szafir DA. Connecting human-robot interaction and data visualization. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, Boulder, USA, March 8–11, 2021, pp. 281–292.

[44] Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 1956, 63(2):81.

[45] Kingma DP, Ba J. Adam: a method for stochastic optimization. In *Proceedings of 3rd International Conference for Learning Representations*, San Diego, USA, May 7–9, 2015.