A Process-Oriented Framework for Robot Imitation Learning in Human-Centered Interactive Tasks

Muhan Hou
Department of Computer Science
Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
m.hou@vu.nl

Koen Hindriks
Department of Computer Science
Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
k.v.hindriks@vu.nl

A.E. Eiben
Department of Computer Science
Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
a.e.eiben@vu.nl

Kim Baraka
Department of Computer Science
Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
k.baraka@vu.nl

***Abstract*— Human-centered interactive robot tasks (e.g., social greetings and cooperative dressing) are a type of task where humans are involved in task dynamics and performance evaluation. Such tasks require spatial and temporal coordination between agents in real-time, tackling physical limitations from constrained robot bodies, and connecting human user experience with concrete learning objectives to inform algorithm design. To solve these challenges, imitation learning has become a popular approach whereby a robot learns to perform a task by imitating how human experts do it (i.e., expert policies). However, previous works tend to isolate the algorithm design from the design of the whole learning pipeline, neglecting its connection with other modules inside the process (like data collection and user-centered subjective evaluation) from the view as a system. Going beyond traditional imitation learning, this work reexamines robot imitation learning in human-centered interactive tasks from the perspective of the whole learning pipeline, ranging from data collection to subjective evaluation. We present a process-oriented framework that consists of a guideline to collect diverse yet representative demonstrations and an interpreter to explain subjective user-centered performance with objective robot-related parameters. We illustrate the steps covered by the framework in a fist-bump greeting task as demonstrative deployment. Results show that our framework is able to identify representative human-centered features to instruct demonstration collection and validate influential robot-centered factors to interpret the gap in subjective performance between the expert policy and the imitator policy.**

***Index Terms*— imitation learning, social greeting, human-centered interactive tasks**

I. Introduction

Human-centered interactive robot tasks are tasks in which humans are involved in the task dynamics and the evaluation of robot performance is highly dependent on human-centered experience. These include for example social greetings [1], cooperative dressing [2], or social navigation [3]. These tasks are challenging since they require spatial and temporal coordination between agents in real-time, tackling physical limitations from constrained robot bodies, and connecting human-centered experience with concrete algorithm-wise learning objectives. Some rule-based methods [4], [5] tried to hardcode interactive responses based on empirical observation, but they can hardly be robust enough for unseen situations and universal for different social agents. Alternatively, data-driven methods [1], [6]–[8] trained the robot with demonstrations from human-human or human-robot interaction, aiming to imitate how human experts perform the tasks (i.e., expert policies). However, these imitation-based approaches mainly focused on improving the algorithm itself and neglect the potential benefits it might bring if more careful attention is paid to other modules inside the design of the learning pipeline, such as data collection and subjective evaluation. This work reexamined robot imitation learning for human-centered interactive tasks from the view of the whole learning pipeline, which we refer to as the "process-oriented" perspective. More specifically, our work aims to tackle two research questions:

- **Q1**: How to collect diverse and representative demonstrations for human-centered interactive tasks?
- **Q2**: How to elicit objective robot-related factors to interpret subjective user evaluations of robot performance and consequently inform algorithm refinement?

For **Q1**, it has been shown that the performance of imitation-based methods relies heavily on the amount and quality of training data [9], [10]. Unfortunately, conditions often only afford limited amounts of data, especially for human-centered interactive tasks where real humans are always required to be present in the loop, which makes a well-designed process for data collection even more important. However, how to collect diverse yet representative demonstration data within the data budget remains an open question. For **Q2**, it is common to evaluate robot performance using subjective metrics [11], [12] in human-centered tasks like social greetings, since it relates more

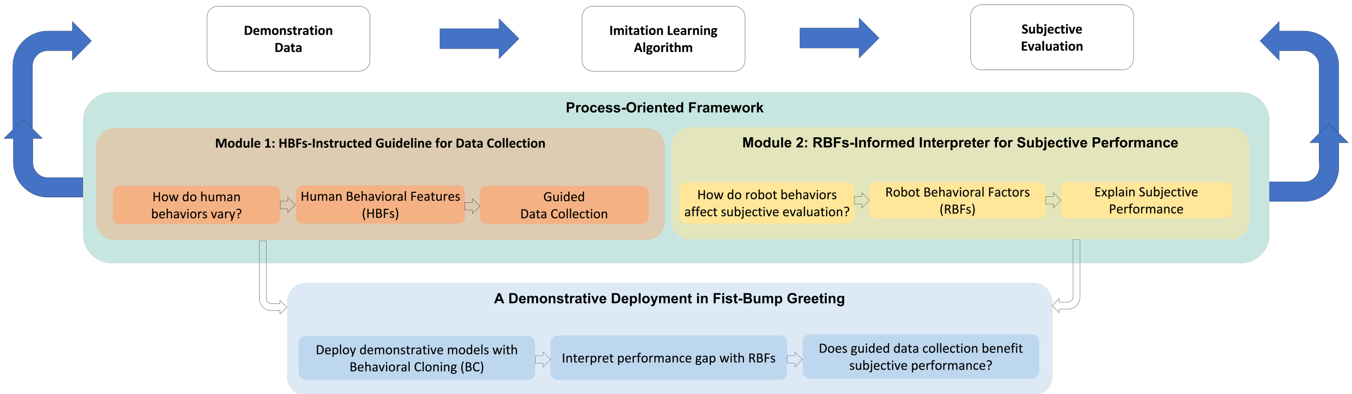

Fig. 1. Overview of our framework and the structure for this paper.

closely to the user experience. However, just knowing how satisfied the human partner feels hardly speaks for how the robot could concretely improve its policy. It would be more improvement-informative if more robot-centered factors that are objective and tuneable could be identified from general subjective metrics. Such objective factors would provide more support to refine learning algorithms to better adapt to user experience, as well as to explain and predict the subjective experience for early intervention. Despite all these benefits, how to identify objective and robot-centered factors from subjective metrics has not yet been fully investigated.

Therefore, we contributed a process-oriented framework (summarized in Fig. 1) to, first, identify human-centered features to inform the collection of diverse and representative demonstrations and, second, validate objective robot-centered factors to explain subjective robot performance. Our contributions can be summarized as follows:

- A guideline to collect diverse and representative demonstration data for human-centered interactive tasks.
- An experimental method to elicit objective robot-centric factors from subjective evaluation to interpret subjective performance gap.
- A demonstrative deployment of our framework in the real-world task of fist-bump greeting.

II. Related Work

A. Robot Learning in Human-Centered Interactive Tasks

Previous work investigated various methods to equip robots with skills for human-centered interactive tasks. Some work employed rule-based methods to control robot behaviors [4], [5]. More advanced methods utilized data-driven methods to imitate human policies from human-human demonstration [1], [6]–[8]. However, few of them tested their methods on an embodied robot to interact with a group of different human agents and retrospected the robot learning from the perspective of the whole learning pipeline (i.e., from data collection to user experience) as opposed to only focusing on algorithm design.

B. Data Collection in Imitation Learning

The performance of learning-based methods is closely related to the quality and quantity of training data [9]. Specifically, data diversity is of great significance to the learning process [10]. To obtain diverse training data, it is common in the field of HRI to collect human data without intervention to capture natural interactions. For instance, researchers of social-aware navigation placed recording devices in public spaces and gathered human trajectories from various perspectives [13]. Similarly, for social greeting tasks, previous work also collected non-verbal behaviors in natural human-human interaction without any exterior instruction [14]. However, in trade for data diversity, such data collection manner often comes with large time and effort cost. Few work investigated into a more structured guideline for HRI data collection that could guarantee data diversity within a constrained data budget. Although [9] presented a workflow for data collection, it did not further prove its functionality in the frame of any learning-based algorithms.

C. Evaluation Metrics in Human-Centered HRI

In the field of HRI, it is common to employ subjective metrics to evaluate robot performance. In terms of social performance, metrics including trust [15], [16], engagement [17], social compliance [18] are often used to describe robot social effectiveness [11]. More systematic metrics like NASA Task Load Index (TLX) [19] and Godspeed likeability [20] are also commonly used to rate user experience from a human-centered perspective. However, results from subjective evaluation are not much informative for algorithm design. Some work attempted to connect subjective evaluation with objective metrics, but they were mainly focused on supplementing [21], [22] or approximating subjective results [12]. By contrast, our work tried to explain the results of human-centered subjective evaluation with robot-centered objective factors and inform algorithm refinement.

III. Methodology

Our framework consists of two experiment-based modules. The first module is the *HBFs-informed guideline for data collection*, which is designed to enhance data quality by facilitating the collection of diverse and representative demonstrations. The second module is the *RBFs-informed* *interpreter for subjective performance*, which aims to interpret human subjective evaluations using objective factors to better inform algorithm refinement.

A. Module 1: HBFs-Informed Guideline for Data Collection

1) Human Behavioral Features (HBFs):
We defined *human behavioral features* as dimensions that could distinguish one person's behavior from another. These features are expected to be measurable and of practical meaning. For example, in the task of fist-bump greeting, an HBF could be the velocity at which the person's hand reaches out. These features can be designed through heuristic observation or feature engineering techniques. However, how to automate and optimize the design of HBFs is beyond the scope of this work.

2) Procedures:
For a given human-centered task and N_{hbf} pre-defined candidate HBFs, each human participant interacted with the robot for $N_{hbf} + 1$ trials, where the robot was controlled by a human expert via the method of Wizard-of-Oz. The whole process consisted of three phases, as shown in Fig. 2. Human participants first intuitively interacted with the robot, and we collected the resultant trajectory ξ_i that consisted of T pairs of (s_t, a_t), i.e., $\xi_i = \{(s_t, a_t)\}_{t=1}^{T}$. s_t was the human-centered state of the participant and a_t was the robot action. For instance, in the task of fist-bump greeting, s_t could be the body pose of the participant and a_t could be the robot joint values. We refer to this phase as *natural data collection*, since no intervention or guidance was provided to human participants during this process. The whole set of demonstrations collected during this phase for all human participants was denoted as $\Gamma_{nat} = \{\xi_i\}_{i=1}^{N_h}$, where N_h was the total number of human participants. After this phase, we provided guidance to human participants, instructing them to vary their original intuitive interactions along the dimension of each candidate HBF. This led to a demonstration set $\Gamma_{hbf}^{k} = \{\xi_i\}_{i=1}^{N_h}$ for each candidate HBF k. We refer to this phase as *guided data collection* and denote the resultant whole set of demonstrations as $\Gamma_{guid} = \{\Gamma_{hbf}^{k}\}_{k=1}^{N_{hbf}}$.

After two phases of data collection, we utilized Feature Selection (FS) to identify representative HBFs from candidate HBFs as valid dimensions to differentiate behaviors of human participants and sorted out the representative demonstrations from Γ_{guide}. We refer to this phase as *representative demonstration curation*. Since the collected demonstrations were essentially trajectories and candidate HBFs were indeed the features to characterize each trajectory, therefore it falls into the category of Unsupervised Feature Selection (UFS) and we chose to employ the classic filter-based UFS method using Laplacian Score (LS) [23] for its independence from learning algorithms and flexible usability. The smaller the Laplacian score L_r of a feature is, the more important the feature is. We used the demonstrations Γ_{nat} from the phase of natural data collection to calculate the Laplacian score for each of the candidate HBFs and chose the top N_{rep} as the representative HBFs. We then only kept the demonstrations Γ_{hbf}^{k} of each representative HBF k in Γ_{guid}, producing the final filtered demonstration set $\Gamma_{guid}^{rep} = \{\Gamma_{hbf}^{k}\}_{k=1}^{N_{rep}}$.

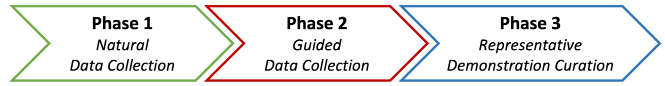

Fig. 2. Procedures of HBFs-informed guideline for data collection.

B. Module 2: RBFs-Informed Interpreter for Subjective Performance

1) Robot Behavioral Factors (RBFs):
We defined the Robot Behavioral Factors (RBFs) as objective and controllable parameters of robot behaviors that may influence subjective evaluation of robot performance in human-centered tasks. For instance, in the task of fist-bump greeting, an RBF can be the reaching velocity of the robot fist. Similar to HBFs, RBFs can be crafted either through task-oriented observation or generated using feature engineering techniques.

2) Procedures:
As shown in Fig. 3, for a given task and predefined N_{rbf} candidate RBFs, each human subject interacted with the robot for N_{rbf} trials, during which the robot was controlled by an expert human teacher via the method of Wizard-of-Oz. During trial i, human participants interacted with the robot twice, once while the robot was in "*normal mode*" and once in "*diversion mode*". In the *normal mode*, human participants intuitively interacted with the "puppet" robot, which followed the exact command from the "performer" robot that was physically controlled by the expert human teacher. By contrast, in the *diversion mode*, we hardcoded the puppet robot to deliberately divert from the original behaviors of the performer robot along the dimension of the candidate RBF i. To enable human participants to detect the distinction between the two modes, we produced the puppet behaviors in diversion mode via a constant divergence ratio $r_{div} = |\phi_{nor}^{i} - \phi_{div}^{i}| \,/\, |\phi_{nor}^{i}|$, where ϕ_{nor}^{i} and ϕ_{div}^{i} represented the value of RBF i in the normal and diversion mode respectively. Upon completion of both modes, every human subject was requested to rate the robot's performance in the task using a 5-point Likert Scale, where 1 indicates "very poor" and 5 indicates "excellent".

Once all N_{rbf} trials were finished, paired samples t-tests would be conducted for each candidate RBF to investigate the influence of the control mode on subjective scores. Candidate RBFs that exhibited significant differences in subjective scores between the two modes would be identified as *influential RBFs*. These influential RBFs served as the output for the module of the RBFs-informed interpreter. When testing imitation learning algorithms for human-centered tasks, these influential RBFs can be used to explain the gap in subjective performance between different baselines. How to utilize them to interpret subjective results will be illustrated in the demonstrative application discussed in section VI.

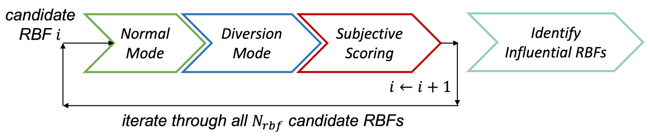

Fig. 3. Procedures of RBFs-informed interpreter

IV. Demonstrative Deployment in the Task of Fist-Bump Greeting

We demonstrated how to deploy our framework in human-centered interactive tasks using the fist-bump greeting as the benchmark task. We chose it to showcase the demonstrative deployment since it is one of the most typical human-centered interactive tasks that can fully reflect the characteristics of such task type: humans play their roles in task dynamics and the evaluation of robot performance is dependent on subjective user experience. Using Behavioral Cloning (BC) as the representative algorithm for imitation learning, we further extended the deployment with an evaluation study, aiming to investigate the potential benefits the module of the HBFs-informed guideline may bring to robot learning and how the module of RBFs-informed interpreter could be utilized to explain the results of subjective performance. All collected demonstrations were available at https://github.com/MH-Hou/process-oriented-framework.git.

A. Demonstrative Deployment of the Framework

1) Task Settings for fist-bump greeting:

In the task of fist-bump greeting, human agents played the role of "initiator" by first reaching out their fists to the robot for bumping. By contrast, the robot responded to the greeting by reaching its fist to meet the human fist and putting it back in an adaptive and timely manner.

We defined the state s_t as a temporal sequence of human body poses for the past w time steps, i.e., $s_t = [P_{t-w+1}, P_{t-w+2}, ..., P_t]$. P_t is the human body poses at time step t and it consists of 3D positions of landmarks including right wrist and right elbow (i.e., $P_t = [p_t^{wrist}, p_t^{elbow}]$). We defined the action a_t as the joint values for the robot right arm (i.e., shoulder, elbow, and wrist) and hip, i.e., $a_t = [q_t^{ShP}, q_t^{ShR}, q_t^{ElY}, q_t^{ElR}, q_t^{WrY}, q_t^{HpR}]$. q_t^{ShP} and q_t^{ShR} represent the pitch and roll of the right shoulder joint, q_t^{ElY} and q_t^{ElR} represent the yaw and roll of the right elbow joint, and q_t^{WrY} represents the right wrist yaw. Similarly, q_t^{HpR} refers to the joint value of hip roll. One complete episode consists of 60 steps, i.e., $T = 60$.

2) Instantiation of HBFs and RBFs:

For the task of fist-bump greeting, we concreted the following 4 candidate HBFs based on heuristic observation:

- **average hand-reaching velocity** (*HBF-1*): It refers to the average velocity of the human hand in the period that the hand has been reaching towards the robot until it stops and waits for the bumping from the robot fist.
- **fist-holding duration** (*HBF-2*): It is defined as the time duration starting from the human agent holding the fist still for bumping to the moment of withdrawing the fist.
- **fist-bumping height** (*HBF-3*): It refers to the height of the human fist in the world frame when it is held still for bumping.
- **average hand-withdrawing velocity** (*HBF-4*): It is defined as the average velocity of the human hand in the period that the human starts to withdraw the fist until it is fully put back and stays still at one side of the body.

For the module of RBFs-informed interpreter, we set the divergence ratio r_{div} as 0.5 and instantiated the following 6 candidate RBFs based on heuristic observation:

- **robot hand-reaching delay** (*RBF-1*): It refers to the time delay between the moment of the human agent reaching out the fist towards the robot and that of the robot reaching out its hand towards the human.
- **robot hand-reaching velocity** (*RBF-2*): It refers to the average velocity of the robot hand during the process it reaches towards the human fist.
- **bumping position offset** (*RBF-3*): It refers to the distance between the human fist and the robot fist when both of them are holding still for the bumping action.
- **robot torso movement** (*RBF-4*): It refers to the movement of the robot's upper body. More specifically, it can be counted as the average deviation of the hip joint value during the whole greeting process.
- **bumping fist orientation** (*RBF-5*): It refers to the fist orientation of the robot hand when it holds still for the bumping action. In our case, it can be calculated as the angle of the robot wrist rotating along the central axis of its forearm.
- **robot hand-withdrawing delay** (*RBF-6*): It refers to the time delay between the moment of the human agent withdrawing the hand and that of the robot withdrawing its hand.

3) Hardware:

We used the humanoid robot Pepper of the SoftBank Robotics company for all experiments, shown in Fig. 4. We used an RGB-camera to capture real-time human body poses and estimated the 3D positions of pose landmarks via MediaPipe [24].

4) Participants:

Following the ethical guidelines by our faculty's research ethics board, we recruited 15 human participants (9 male and 6 female) from campus via poster advertisement. 11 of them were aged between 18 to 29 and 4 of them were aged between 30 to 39. In terms of the experience of interacting with a robot, 3 of them had no previous experience, 10 of them indicated some experience, and 2 of them had extensive experience. After the experiments finished, participants were compensated with a digital gift card of €10 for participation.

B. Evaluation Study

1) BC Model Training:

The structure of the BC model was illustrated in Fig. 5. We optimized BC models by minimizing the loss function

Fig. 4. Demonstrative deployment in the benchmark task of fist-bump.

$\mathcal{L}$ constructed by the cross entropy and expressed as:

$$\mathcal{L} = -\frac{1}{N_s}\sum_{i=1}^{N_s} log(\tilde{P}(a_i^{ex}|s_i^{ex})) \tag{1}$$

where N_s represents the total number of samples in demonstration data and $\tilde{P}(a_i^{ex}|s_i^{ex})$ represents the probability of expert action a_i^{ex} under the expert state s_i^{ex}.

We trained each BC model using the Adam optimizer with a constant learning rate of 0.001 and L2-norm regularization with a regularization rate lambda of 0.001. We trained each BC model for 500 epochs with a batch size of 60.

2) Participants:
Following the same ethical guideline as in section IV-A, we invited another group of 16 human participants (10 male, 5 female, and 1 other gender) from campus via poster advertisement. 2 of them were aged 17 or under, 11 of them were aged between 18 to 29, 1 of them was aged between 30 to 39, and 2 of them were aged between 40 to 49. Regarding the experience of interacting with a robot, 7 of them had no previous experience, 6 of them indicated some experience, and 3 of them had extensive experience. Participation was compensated with a digital gift card of €10.

3) Procedures:

We trained two BC models respectively with the demonstrations collected in the phase of natural data collection (i.e., Γ_{nat}) and guided data collection (i.e., Γ_{guid}^{rep}). For a fair comparison, we randomly sampled from the filtered demonstration set Γ_{guid}^{rep} to obtain an equal amount of demonstrations as Γ_{nat}. For convenience, we denoted the resultant BC models as *natural BC model* and *guided BC model* respectively.

To evaluate the subjective performance of the guided BC model, we conducted an evaluation study that consisted of 3 sections of experiments. In each section, the robot was controlled by a different policy (i.e. the guided BC model, or Wizard-of-Oz). Human participants conducted fist-bump with the robot in their most intuitive ways and scored robot performance via a 5-point Likert Scale where 1 represents "very poor" and 5 represents "excellent". A 1-minute break followed after human participants finished evaluating and then we repeated this "greet-evaluate" procedure until all sections were completed. For the concern of the carryover effect, we randomly selected the section order for each human subject.

To further investigate the potential benefits the module of the HBFs-informed guideline could bring to robot learning, we also compared the imitation performance of two trained BC models. Using the Wizard-of-Oz demonstrations collected during the evaluation experiments as the testing expert demonstrations, we evaluated how closely the trained BC model resembled an expert demonstration ξ^e via the imitation error e_{imit} defined as:

$$e_{imit} = \frac{1}{T}\sum_{(a_t^e, s_t^e)\in\xi^e} \|\pi^{BC}(s_t^e) - a_t^e\| \tag{2}$$

where T is the total number of state-action pairs in the expert demonstration ξ_e and π^{BC} is the trained BC policy.

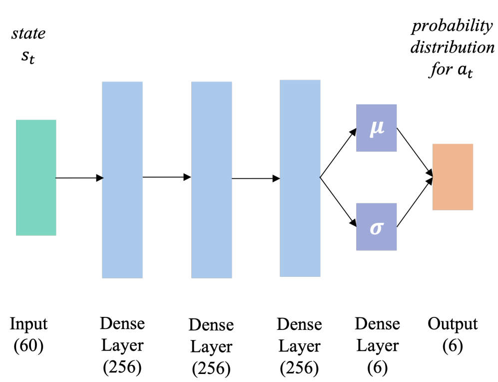

Fig. 5. The Neural Network architecture for our Behavioral Cloning model. The input is the state s_t defined as the 3D body poses of the human for the past w time steps. We used $w = 10$ and 6 landmarks, leading to an input dimension of 60. μ and σ represent the mean and standard deviation of the probability distribution of output action a_t.

V. RESULTS

A. Results for HBFs-Informed Guideline

After calculating the candidate HBFs for each demonstration in Γnat and applying feature selection to each candidate HBF, we obtained their corresponding Laplacian Scores shown in Table I. Since a lower Laplacian Score indicates a more important feature, the results showed that human behaviors in fist-bump greeting were mostly different from each other in the aspect of average hand-withdrawing velocity (i.e., HBF-4), followed by average hand-reaching velocity (i.e., HBF-1), fist-holding duration (i.e., HBF-2), and fist-bumping height (i.e, HBF-3). Furthermore, since there was a large gap between the Laplacian Score of HBF-2 and HBF-3, it indicated that human behaviors in fist-bump greeting were far less different regarding the fist-bumping height, as compared with other potential HBFs. Therefore, we ruled out the HBF-3 and considered HBF-4, HBF-1, and HBF-2 as the *representative HBFs* (i.e., $N_{rep} = 3$).

TABLE I
FEATURE SELECTION RESULTS FOR CANDIDATE HBFS

	HBF-1	HBF-2	HBF-3	HBF-4
Laplacian Score	**0.111**	**0.248**	0.649	**0.0827**

B. Results for RBFs-Informed Interpreter

For each candidate RBF, we conducted a paired samples t-test to investigate the influence of the control mode on subjective scores, shown in Fig. 6. More specifically:

- RBF-1 (robot hand-reaching delay): There was a significant difference in subjective scores between normal mode ($M = 3.73, SD = 0.77$) and diversion mode ($M = 2.46, SD = 0.88$); $t(14) = 3.68,\ p < .05$ with a large effect size (Cohen's $d = 1.47$).
- RBF-2 (robot hand-reaching velocity): There was a significant difference in subjective scores between normal mode ($M = 3.73, SD = 0.68$) and diversion mode ($M = 2.53, SD = 0.81$); $t(14) = 4.94,\ p < .001$ with a large effect size (Cohen's $d = 1.56$).
- RBF-3 (bumping position offset): There was a significant difference in subjective scores between normal mode ($M = 3.33, SD = 1.07$) and diversion mode ($M = 2.53, SD = 0.81$); $t(14) = 2.70,\ p < .05$ with a large effect size (Cohen's $d = 0.81$).
- RBF-4 (robot torso movement): There was a marginally significant difference in subjective scores between normal mode ($M = 3.87, SD = 0.72$) and diversion mode ($M = 3.2, SD = 1.17$); $t(14) = 2.00,\ p = .07$ with a medium effect size (Cohen's $d = 0.67$).
- RBF-5 (bumping fist orientation): There was no significant difference in subjective scores between normal mode ($M = 3.53, SD = 0.96$) and diversion mode ($M = 3.60, SD = 0.71$); $t(14) = -0.19,\ p = .85$ with a small effect size (Cohen's $d = 0.08$).
- RBF-6 (robot hand-withdrawing delay): There was a significant difference in subjective scores between normal mode ($M = 4.13, SD = 0.81$) and diversion mode ($M = 3.20, SD = 0.98$); $t(14) = 3.76,\ p < .05$ with a large effect size (Cohen's $d = 1.01$).

Since we found significant differences for RBF-1, RBF-2, RBF-3, RBF-4, and RBF-6, we identified these candidate RBFs as the *influential RBFs*, meaning they were able to significantly influence subjective evaluation of robot performance in the task of fist-bump greeting.

C. Results for Evaluation Study

To compare the subjective performance between the guided BC model and Wizard-of-Oz, we conducted a paired samples t-test to investigate the influence of the type of control policies on subjective scores. Results indicated that there was no significant difference in subjective scores between Wizard-of-Oz ($M = 4.125, SD = 1.05$) and the guided BC model ($M = 3.75, SD = 1.03$); $t(15) = 1.031,\ p = .32$ with a small effect size (Cohen's $d = 0.35$).

To explain the results of subjective performance with influential RBFs, we also conducted a paired samples t-test for each influential RBF to investigate the influence of the type of control policies on the influential RBF, shown in Fig. 7. More specifically:

- RBF-1: There was no significant difference in the RBF value between Wizard-of-Oz ($M = 0.153, SD = 0.161$) and the guided BC model ($M = 0.263, SD = 0.242$); $t(15) = -1.380,\ p = .19$ with a medium effect size (Cohen's $d = 0.52$).
- RBF-2: There was a significant difference in the RBF value between Wizard-of-Oz ($M = 0.457, SD = 0.248$) and the guided BC model ($M = 0.227, SD = 0.230$); $t(15) = 4.211,\ p < .001$ with a large effect size (Cohen's $d = 0.93$).
- RBF-3: There was no significant difference in the RBF value between Wizard-of-Oz ($M = 0.594, SD = 0.231$) and the guided BC model ($M = 0.600, SD = 0.196$); $t(15) = -0.093,\ p = .93$ with a small effect size (Cohen's $d = 0.03$).
- RBF-4: There was a significant difference in the RBF value between Wizard-of-Oz ($M = 0.311, SD = 0.154$) and the guided BC model ($M = 0.726, SD = 0.126$); $t(15) = -9.130,\ p < .001$ with a large effect size (Cohen's $d = 2.85$).
- RBF-6: There was a significant difference in the RBF value between Wizard-of-Oz ($M = 0.168, SD = 0.271$) and the guided BC model ($M = 0.490, SD = 0.252$); $t(15) = -3.51,\ p < .05$ with a large effect size (Cohen's $d = 1.19$).

To further explore the potential benefits that the module of the HBFs-informed guideline may bring to robot learning, we conducted a paired samples t-test to investigate the influence of the type of BC models on the imitation error e_{imit}. We observed no significant difference in the imitation error between the natural BC model ($M = 0.444, SD = 0.159$) and the guided BC model ($M = 0.492, SD = 0.123$); $t(15) = -1.516,\ p = .15$ with a small effect size (Cohen's $d = 0.33$).

VI. Discussion

Our results from the demonstrative deployment and evaluation study indicated that our framework was able to identify representative HBFs for collecting diverse and representative demonstrations. Furthermore, it could also elicit influential RBFs to interpret the user-centered subjective evaluation with objective robot-related parameters and utilize it to inform algorithm refinement.

More specifically, for the results of the HBFs-informed guideline, we found that human behavior in fist-bumping greeting are most different from each other in aspects of average hand-withdrawing velocity (i.e., HBF-4), average hand-reaching velocity (i.e., HBF-1), and fist-holding duration (i.e., HBF-2). Although it was true that human heights were diverse from each other and expected to be reflected in their greeting interaction with the robot, they also tended to adapt their greeting to that of the robot in such interactive and cooperative task. In our case, the Pepper robot we used was with a height of about 1.2m and an arm length of about 0.6m. This child-like body size was much smaller than that of our human participants, making it more common for human participants to adapt their greeting trajectories to the space that the robot could reach. Therefore, it was reasonable that

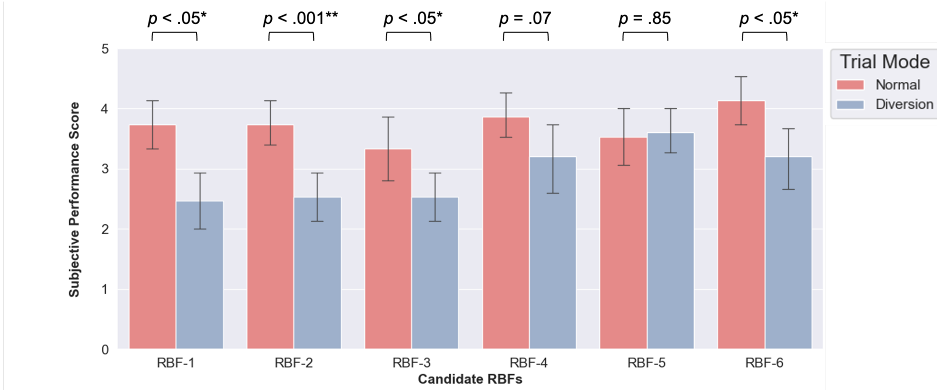

Fig. 6. Results for subjective performance scores in the experiments of candidate RBFs in demonstrative deployment.

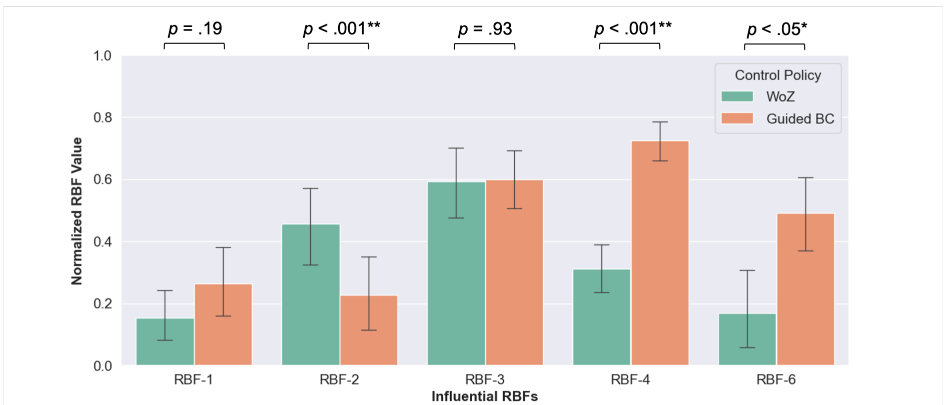

Fig. 7. Results for normalized values of influential RBFs in the experiments of the evaluation study.

human greeting behaviors varied less in fist-bumping height (i.e., HBF-3) than other HBFs.

For the results of the RBFs-informed interpreter, we found that robot hand-reaching delay (RBF-1), robot hand-reaching velocity (RBF-2), bumping position offset (RBF-3), robot torse movement (RBF-4), and robot hand-withdrawing delay (RBF-6) were able to influence the subjective evaluation of robot performance. This indicated that these influential RBFs could be utilized to interpret the subjective results from objective perspectives and might be able to better inform the algorithm refinement if there exists a gap in subjective performance between different baselines.

For the subjective performance of the guided BC model, the results showed that it was not significantly different from that of the method of Wizard-of-Oz. Considering that the Wizard-of-Oz method actually represented the expert policy that the BC model aimed to imitate, such a result indicated a potentially positive effect of our HBFs-informed guideline on subjective performance. However, further confirmation is necessary through larger sample sizes and more extensive experiments.

Furthermore, we found that, among all the influential RBFs, robot hand-reaching velocity (RBF-2), robot torso movement (RBF-4), and robot hand-withdrawing delay (RBF-6) were significantly different between the Wizard-of-Oz method and the guided BC model. More specifically, the guided BC model exhibited a slower robot hand-reaching velocity, larger robot hand-withdrawing delay, and larger torso movements, as compared with the Wizard-of-Oz method. While a slower reaching velocity and larger withdrawing delay lead to a less timely fist-bump, larger torso movements make the robot more lively and adorable. Indeed, some human participants made positive comments about the robot employing the whole torso to do the fist-bump greeting, as opposed to only using the right arm joints. The differences in these three influential RBFs compensated for each other, leading to an insignificant gap in subjective performance between the Wizard-of-Oz and the guided BC model. If more attention could be paid to the reaching velocity and withdrawing delay (e.g., reweighting them in the cost function), it might lead to further improvement of robot performance in subjective evaluation.

For the potential learning benefits of the HBFs-informed guideline, we found no significant improvement in imitation performance between the guided BC model and the natural BC model. One possible reason for this could be the limited amount of demonstrations collected, given the data-intensive nature of imitation learning algorithms. However, we believe that with a larger number of demonstrations and testing experiments conducted on a larger sample size, more noticeable improvements can be achieved.

A. Limitations and Future Work

For the perception part of our work, we used a RGB-camera without depth information. Although we fixated the distance between the robot base and the standing position of human participants during all the experiments, it would yet cause inaccuracy in body pose estimation and impact the quality of data collection. We plan to employ RBG-D camera in the future to acquire more accurate perception of human movement and allow for more flexible interaction between humans and robots.

Also, we only indicated the possible correlation between RBFs and subjective evaluation of robot performance without further investigating how exactly these two were correlated with each other. With further study in the future, it may bring more concrete insights for the algorithm design of imitation-based learning to perform better in subjective evaluation.

Lastly, we only built our studies around the task of fist-bump as one of the most typical greeting behaviors. The results may differ when it comes to other types of greeting (e.g., handshaking and high-five) or even different types of coordination-involved tasks (e.g., social-aware navigation). In the future, we also plan to apply the results from this work to a broader range of cooperative task settings and further investigate the potential of imitation learning algorithms in social interactive scenarios.

VII. Conclusion

In this work, we presented a process-oriented framework of robot imitation learning for human-centered interactive tasks. Different from previous works that were predominantly algorithm-centric, our framework reexamined robot imitation learning from the perspective of the whole process from data collection to subjective evaluation. More specifically, it consisted of a guideline to collect higher quality demonstrations that were diverse and representative along important human-centered features (i.e., HBFs) and an interpreter to explain the subjective results with robot-centered objective parameters (i.e., RBFs). Furthermore, we provided a demonstrative deployment of our framework in the task of fist-bump greeting. Results indicated that human fist-bump greetings were significantly diverse and distinct in terms of hand-reaching velocity, fist-holding duration, and hand-withdrawing velocity, which were able to be utilized to instruct demonstration collection. Also, results showed that robot subjective performance was influenced by hand-reaching delay, hand-reaching velocity, bumping position offset, torso movement, and hand-withdrawing delay. Furthermore, the gap in subjective performance between the expert policy and learned policy was able to be explained in terms of robot hand-reaching velocity and torso movement.

References

[1] J. Bütepage, A. Ghadirzadeh, Ö. Öztimur Karadağ, M. Björkman, and D. Kragic, "Imitating by generating: Deep generative models for imitation of interactive tasks," *Frontiers in Robotics and AI*, vol. 7, p. 47, 2020.

[2] J. Zhu, M. Gienger, G. Franzese, and J. Kober, "Do you need a hand?– an interactive robotic dressing assistance scheme," *arXiv preprint arXiv:2301.02749*, 2023.

[3] Y. Che, A. M. Okamura, and D. Sadigh, "Efficient and trustworthy social navigation via explicit and implicit robot–human communication," *IEEE Transactions on Robotics*, vol. 36, no. 3, pp. 692–707, 2020.

[4] A. Zaraki, L. Wood, O. Novanda, B. Robins, and K. Dautenhahn, "Toward autonomous child-robot interaction: development of an interactive architecture for the humanoid kaspar robot," 2017.

[5] E. Okamura and F. Tanaka, "Design of a robot that is capable of high fiving with humans," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 704–711, IEEE, 2017.

[6] H. B. Amor, D. Vogt, M. Ewerton, E. Berger, B. Jung, and J. Peters, "Learning responsive robot behavior by imitation," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3257–3264, IEEE, 2013.

[7] H. B. Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters, "Interaction primitives for human-robot cooperation tasks," in *2014 IEEE international conference on robotics and automation (ICRA)*, pp. 2831–2837, IEEE, 2014.

[8] W.-R. Ko, J. Lee, M. Jang, and J. Kim, "End-to-end learning of social behaviors for humanoid robots," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 1200–1205, IEEE, 2020.

[9] Z. Han and T. Williams, "Towards formalizing hri data collection processes," *arXiv preprint arXiv:2203.08396*, 2022.

[10] Z. Gong, P. Zhong, and W. Hu, "Diversity in machine learning," *IEEE Access*, vol. 7, pp. 64323–64350, 2019.

[11] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common metrics for human-robot interaction," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pp. 33–40, 2006.

[12] Y. Mizuchi and T. Inamura, "Optimization of criterion for objective evaluation of hri performance that approximates subjective evaluation: a case study in robot competition," *Advanced Robotics*, vol. 34, no. 3-4, pp. 142–156, 2020.

[13] A. Rudenko, T. P. Kucner, C. S. Swaminathan, R. T. Chadalavada, K. O. Arras, and A. J. Lilienthal, "Thör: Human-robot navigation data collection and accurate motion trajectories dataset," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 676–682, 2020.

[14] W.-R. Ko, M. Jang, J. Lee, and J. Kim, "Air-act2act: Human–human interaction dataset for teaching non-verbal social behaviors to robots," *The International Journal of Robotics Research*, vol. 40, no. 4-5, pp. 691–697, 2021.

[15] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human factors*, vol. 46, no. 1, pp. 50–80, 2004.

[16] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, "Foundations for an empirically determined scale of trust in automated systems," *International journal of cognitive ergonomics*, vol. 4, no. 1, pp. 53–71, 2000.

[17] A. Bruce, I. Nourbakhsh, and R. Simmons, "The role of expressiveness and attention in human-robot interaction," in *Proceedings 2002 IEEE international conference on robotics and automation (Cat. No. 02CH37292)*, vol. 4, pp. 4138–4142, IEEE, 2002.

[18] J. Goetz and S. Kiesler, "Cooperation with a robotic assistant," in *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, pp. 578–579, 2002.

[19] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*, vol. 52, pp. 139–183, Elsevier, 1988.

[20] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International journal of social robotics*, vol. 1, no. 1, pp. 71–81, 2009.

[21] F. Dehais, E. A. Sisbot, R. Alami, and M. Causse, "Physiological and subjective evaluation of a human–robot object hand-over task," *Applied ergonomics*, vol. 42, no. 6, pp. 785–791, 2011.

[22] G. Hoffman, "Evaluating fluency in human–robot collaboration," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 3, pp. 209–218, 2019.

[23] X. He, D. Cai, and P. Niyogi, "Laplacian score for feature selection," *Advances in neural information processing systems*, vol. 18, 2005.

[24] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, *et al.*, "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.