



Investigating the Effects of Robot Engagement Communication on Learning from Demonstration

Mingfei Sun¹ · Zhenhui Peng¹ · Meng Xia¹ · Xiaojuan Ma¹

Accepted: 30 August 2021 / Published online: 8 October 2021
© The Author(s) 2021

Abstract

Robot learning from demonstration (RLfD) is a technique for robots to derive policies from instructors' examples. Although the reciprocal effects of student engagement on teacher behavior are widely recognized in the educational community, it is unclear whether the same phenomenon holds for RLfD. To fill this gap, we first design three types of robot engagement behavior (gaze, imitation, and a hybrid of the two) based on the learning literature. We then conduct, in a simulation environment, a within-subject user study to investigate the impact of different robot engagement cues on humans compared to a “without-engagement” condition. Results suggest that engagement communication has significantly negative influences on the human's estimation of the simulated robots' capability and significantly raises their expectation towards the learning outcomes, even though we do not run actual imitation learning algorithms in the experiments. Moreover, imitation behavior affects humans more than gaze does in all metrics, while their combination has the most profound influences on humans. We also find that communicating engagement via imitation or the combined behavior significantly improves humans' perception towards the quality of simulated demonstrations, even if all demonstrations are of the same quality.

Keywords Robot communicating engagement · Robot learning from demonstrations · Robot behavior in learning from demonstration · Robot simulation

1 Introduction

Robot learning from demonstration (RLfD) is a technique where a robot derives a mapping from states to actions, a.k.a *policy*, from instructors' demonstrations [7]. This technique is successful in teaching robots physical skills by imitating instructors' body movements e.g., pole balancing [7], tennis swings [31], air hockey maneuvers [9], etc. A standard RLfD process takes two steps: demonstration gathering step, which collects demonstrations from the human demonstrators, and policy deriving step, which reasons the underlying state-action mappings [4]. Like a human learner, a robot in RLfD could have different strategies of gathering demonstrations according to its underlying policy derivation algorithms. For example, robots with the DAgger algorithm

[50] learn progressively by taking incremental demonstrations from instructors, much like going through a scaffolding process [32,51]. A robot can also learn more proactively. For example, if equipped with Confidence-Based Autonomy (CBA) [17], an interactive algorithm for RLfD, a robot can request demonstrations at the states of which it has little or no knowledge. These learning strategies have been proven to be very effective and thus widely adopted in RLfD [40].

Unlike human learners, robots in previous RLfD processes rarely show any engagement cues during the learning process. They mostly remain stationary without giving any feedback, especially when instructors are giving demonstrations (i.e., in the demonstration gathering step). In human tutelage, engagement cues play an important role in shaping instructors' mental model of the learners [58]. For example, learners' attentional engagement, e.g., gaze, indicates their points of interest in the instructions. Attention direction is one of the essential mechanisms that contributes to the learning process. On the other hand, imitation, a gesture engagement cue, shows learners' motivation to perform like the instructors [16]. It is reported that learner engagement cues could potentially affect instructor perceptions and behavior [25].

✉ Mingfei Sun
mingfei.sun@alumni.ust.hk

Xiaojuan Ma
mxj@cse.ust.hk

¹ Department of CSE, Hong Kong University of Science and Technology, Hong Kong SAR, China

For example, in educational research, instructors are found to tend to provide more support to learners of high gesture engagement [58]. The gaze and gesture engagement are also reported to be specific to a learning process [16] while other cues are common to any interactions.

These effects of showing these learning-specific engagement cues, however, are less explored in the RLfD research, partly because designing engagement cues for robots in the context of RLfD is challenging. First, most of the existing methods for generating engagement cues in human–robot interaction (HRI) cannot be directly applied to RLfD. For example, it is common practice in HRI to simulate robots' attentional engagement by directing their gaze towards visually salient elements (e.g., color or lightness [45]), specific objects (e.g., human faces [54]) or predefined events (e.g., pointing gestures [11]). This practice cannot be easily set up in RLfD because the robot's allocation of attention should follow the development of instructors' demonstrations. This is especially true in skill-oriented RLfD, where the robot needs to reproduce the body skills from the human demonstrator. In this context, the attention should be subject to the demonstrations, i.e., body movements, which are less constrained and highly dynamic compared to a standard HRI process. Methods for generating other engagement cues, e.g., imitation [8,41,49], also need further adaptation to accommodate the dynamic nature of RLfD, especially when the human body and the robot structures are not aligned (a.k.a. the correspondence problem [4]). Second, even if an engagement cue can be designed effectively, its deployment in RLfD should be in real-time with a low computational cost.

To adapt and improve existing attention and imitation methods to be used in RLfD, we focus on skill-oriented RLfD and propose two novel methods (*Instant attention* and *Approximate imitation*) to enable robots to communicate their learning engagement in an RLfD process. Note that we consider the demonstration gathering step as the interaction scenario since it determines the demonstration quality, which is crucial for the policy optimality [4,61]. We do not focus on designing effective learning algorithms for demonstration learning. The learning engagement cues are generated as follows: the *Instant attention* method generates robot gaze engagement by tracking instructors' body movements through particle filters; the *Approximate imitation* method produces gesture engagement, i.e., imitation, by partially mapping the instructor's joint movements to those of the robot with approximations. We then use the proposed methods to generate three modes of engagement communication (via attention, via imitation, and a hybrid of the two) for robots in RLfD. To investigate the effects of the three engagement modes on humans, we compare them with another mode ("without-engagement" in which

the robot remains stationary¹ as most robots do in existing RLfD studies [7,9,31]) by a within-subject user study in a simulation environment. Results suggest that robots with the proposed cues are perceived to be more engaged in the learning process and their behaviors are more socially acceptable in RLfD than the robots without. Also, having engagement cues significantly affects human's estimation of the robots' learning capabilities, making their estimation over-optimistic. The robots which communicate engagement in RLfD are perceived to be significantly more capable in learning than the robots without, even though none of them are equipped with imitation learning algorithms.² Engagement communication also affects the human's expectation towards the final learning outcomes. Furthermore, gesture cues influence humans' perceptions significantly more than gaze engagement does, while the hybrid cues significantly outperform the other two. We also find that showing gesture or combined engagement significantly improves humans' evaluation of demonstration quality. Specifically, the human participants perceived the demonstrations to be significantly more appropriate for the robot to learn when the robot communicates its engagement via gesture or mixed engagement, even though all demonstrations are of the same quality.

The contributions of this paper are as follows. First, we propose two novel algorithms which allow robots to generate attention and imitation behavior to communicate their learning engagement with low computations in RLfD. Sec-

¹ Note that in some types of RLfD the robot does not remain stationary, e.g., kinesthetic teaching, teleoperation and collaborative tasks. For these tasks, the robot thus may need other types of engagement cues, which is out of scope of this paper.

² When referring to learning algorithms, we mean the classic and state-of-the-art imitation learning methods, e.g., inverse reinforcement learning [46], generative adversarial imitation learning [28,61]. Note that the engagement generation process is fundamentally different from a learning process in RLfD. Engagement generation relies on the interaction process itself. Without the interaction, engagement means nothing. We thus design the algorithms for robots to actively involve themselves in the interaction process, paying attention to the demonstrator's joints and behaving approximately as the demonstrator's moves. The robot with these two algorithms needs the demonstrator's full participation, otherwise, no engagement communication could then be triggered. Such engagement generation actively relies on the demonstrator's moves. However, the learning process in any learning from demonstration aims to find an autonomous policy, with which the robot could reproduce the demonstrated moves autonomously without any demonstration being replayed. In other words, the goal of learning is to remove such reliance on demonstrators. For example, in any modern learning from demonstration algorithms, behavioral cloning, or inverse reinforcement learning, the human demonstrations are often used as a proxy to define the loss or the reward. In addition, though the training is to derive a policy, a mapping from states to actions, the states and actions are defined differently in RLfD. For example, the robot state in RLfD is often defined as demonstration-agnostic or robot-centric so that the learned policy could then be triggered without any demonstrators being involved.

ond, we developed a simulation platform to evaluate the effect of engagement communication in RLfD. Third, we take a first step towards evaluating the effects of three types of engagement cues (attention, imitation, and hybrid) on humans. Through evaluation in a simulation environment with a humanoid robot learning the different skills from a simulated demonstrator, we show interesting findings on the design of robot engagement communication in RLfD. To the best of our knowledge, this paper is the first to systematically investigate how robot engagement communication affects the humans' perceptions and expectations of the robot in RLfD.

2 Related Work

2.1 Robot Learning from Demonstration (RLfD)

Robot Learning from Demonstration (RLfD) is also known as “Programming by Demonstration”, “imitation learning”, or “teaching by showing” [52]. Rather than exhaustively searching the entire policy space, RLfD enables robots to derive an optimal policy from demonstrators' (also called instructors) demonstrations [7]. Usually, this technique does not require additional knowledge about programming and machine learning from human instructors, and thus opens up new possibilities for common users to teach robots [18]. Existing studies on RLfD focus mainly on policy derivation algorithms, e.g., mapping states to actions by supervised learning [17], updating the policy by value iteration in Reinforcement Learning [7], and recovering rewards to explain demonstrations by Inverse Reinforcement Learning [1,61]. Some studies also work on designing robots' reciprocal learning feedback to communicate what the robots have learned to human teachers, e.g., demonstrating the robot's current learned policy [12], providing verbal and/or nonverbal cues [2,11,13,36,48,63], or visualizing where they succeed and fail [53]. These studies, however, largely overlook how the robots' engagement behavior would affect the instructors and their demonstrations, especially during the demonstration gathering step. Hence, in this work, we consider how to generate behavior that allows robots to communicate their learning engagement to instructors, and investigate their potential effects on RLfD.

2.2 Engagement and Learning Engagement Cues

Engagement is a broad concept in HRI with many different definitions. Some studies focus on the whole spectrum of interaction, and defines engagement as the process of initiating, maintaining, and terminating the interaction between humans and robots [55]. Others narrow the notion of engagement down to the maintenance of interactions, interpreting

engagement as humans' willingness to stay in the interaction [65,69].

In the context of learning, engagement mainly refers to the state of being connected in the learning interaction, which can be measured from three aspects: cognition, behavior, and emotion [56]. Cognitive engagement is closely related to the allocation of attention as it is one of the most important cognitive resources [47]. Failure to attend to another person indicates a lack of interest [5]. Thus, we adopt attention as a cue to communicate cognitive engagement in RLfD. Gesture engagement is captured by task-related behavior, e.g., task attempts, efforts, active feedback, etc. Imitation, a common gesture engagement signal, refers to “non-conscious mimicry of the postures, mannerisms, facial expressions, (speech), and other behaviors of one's interaction partners” [14]. In interpersonal communications and HRI, the imitation behavior increases the likelihood of understanding [15], interpersonal coordination [10] and emotional contagion [27]. In the context of learning, the imitation behavior also indicates the robot's internal status in learning, e.g., the progress and motivation [16]. Thus, we use imitation as a way to communicate the gesture engagement for robots in RLfD. Emotional engagement is associated with the affective states evoked by the interaction, including valence and arousal. Despite its importance, emotional engagement is hard to apply in RLfD since most existing RLfD robot systems lack the full ability to express emotions. In the scope of this paper, we define the robot learning engagement as the involvement in the learning process, with a focus on its cognitive engagement, i.e., attention, and gesture engagement, i.e., imitation.³ The following subsection presents related work on generating attention and imitation behavior to communicate engagement.

2.3 Robots' Communication of Engagement

In HRI, a robot can communicate its **attention** via different physical channels, e.g., gaze [38,42–44], head orientation [42,63], and body postures [66]. Regardless of which channel they use, robots are usually programmed to pay attention to salient elements, including but not limited to colors [11], objects with visual intensity [45], and movements [11,45]. For example, Nagai et al. regarded visually outstanding points in the surroundings, in terms of their colors, intensities, orientations, brightness, and movements, as points of attention [45]. Other work directs robots' attention to

³ Note that one survey on RLfD [4] describes many types of modeling approaches and methods of acquiring demonstrations. This paper is primarily concerned with getting demonstrations with visual input, and for other ways like kinesthetic teaching or teleoperation, engagement cues may need to be defined differently, if possible. In addition, other types of tasks, such as collaborative ones between two robots or between a robot and a human, may also require different engagement cues.

specific objects, e.g., human faces [54] and colorful balls [3] to name a few, or predefined events, e.g., pointing gestures [42]. For example, Sidner et al. designed a robot that pays attention to participants' faces most of the time [54]. Lockerd et al. drove the robot attention mechanism with interaction events, such as looking at an object when it is pointed at or looking at a subject when the person takes a conversational turn [42]. To accommodate multiple events, a state transition diagram is usually adopted to control any attention shifts [11,42]. Though these studies provide insightful information about the design of robot attention, their approaches may not easily apply to skill-oriented RLfD as the point for attention in instructors' body movements is dynamically changing. In an RLfD process, the robot may be required to learn how to manipulate a specific object or to perform a specific action, e.g., walking like a human being. It may be true in the first case that the engagement could be communicated by paying attention to the object when it is being manipulated. However, in the second case of learning an action or a skill, it is hard to define the salient elements as the whole demonstration motion would be of interest.

Compared to attention, the *imitation* behavior has been less widely adopted as a robot engagement cue. The robot imitation of a human participant's behavior in real-time is inherently challenging due to the correspondence problems [4] as well as the robot's physical constraints [34,35,60]. Hence, instead of generating full-body imitation behavior, some HRI researchers proposed to do partial imitations. For example, Baileson and Yee built an immersive virtual agent that subtly mimicked people's head movements in real-time [8]. A similar imitation strategy was applied by Riek et al. to a chimpanzee robot [49]. In addition to head imitation, gesture "mirroring" has also been implemented by Li et al. on a robot confederate [41]. Although these studies showed that partial imitation behavior improves participants' perception of robots' capabilities [21,23], they mainly used ruled-based methods [8] or predefined behavior [41], which may not be transferable to RLfD scenarios. In this work, we employ the same strategy and allow robot learners to make partial imitations. Different from existing work, we take an algorithmic approach to automatically generating approximate imitations of instructors' body movements for robots in real-time.

3 Learning Engagement Modeling

This section presents two methods for generating engagement cues. The first subsection briefly introduces human body poses and forms the basis of the proposed methods. The remaining subsections describe the methods in detail.

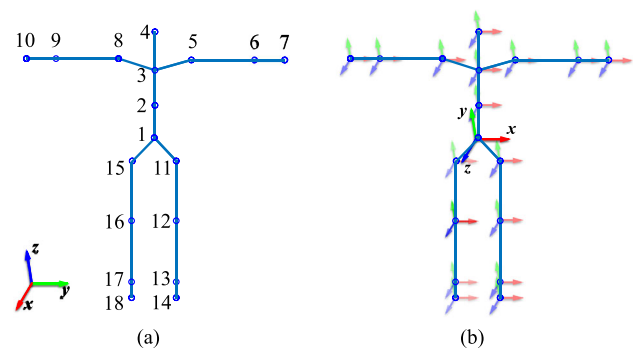


Fig. 1 **a** A body pose in the *position form*: all joints are described in a single frame by their positions; **b** A body pose in the *transformation form*: each joint has its frame and the skeleton defines parent–child structures and translations between frames; the frame with x – y – z labels is the root frame and is referred in the sensor frame

3.1 Representation of the Body Pose

In RLfD, instructors usually demonstrate a point via their body movements. Our proposed methods thus use human body poses to generate attention and imitation behavior. A body pose is usually depicted by a tree-like skeleton, with nodes as human joints and links as body bones (shown in Fig. 1). Mathematically, this skeletal structure can be represented in two forms⁴: the *position form* and the *transformation form*.

Position form The position form describes the body pose in a single frame of reference (usually the sensor frame), as shown in Fig. 1a. In this form, the pose skeleton is denoted as $[J^{(1)}, J^{(2)}, \dots, J^{(n)}]$, where $J^{(i)} \in \mathbb{R}^3$ is the position vector of the i -th joint in the skeleton, and n is the number of joints. This form gives each joint its global position, providing the potential attention point for the robot. Hence it is used for the *Instant attention* algorithm to generate robot attention points.

Transformation form The transformation form describes the body pose in a series of frames of reference [62], as shown in Fig. 1b. In particular, each joint has its frame (a right-handed frame), and the links in the tree-like skeleton define parent–child structures between frames. The pose of a non-root joint is then described by a translation (i.e., the bone length) and a rotation (i.e., joint movement) in its parental frame, with the root joint (often the hip joint) described in the sensor frame. This form decomposes a human body movement into joint rotations (body-independent) and joint translations (body-dependent) in a way that the movement can be easily imitated by robots: just mapping the rotations onto robot joints. We denote this form as $[T_1, T_2, T_3, \dots, T_n]$, and use it for the *Approximate imitation* algorithm to obtain approximate imitation behavior.

⁴ Usually, the two forms are readily available in most body pose extracting sensors, e.g., Kinect.

3.2 Instant Attention

The gaze engagement for robots is generated based on the cognitive theories on human attention. Generally speaking, a generation process of human visual attention involves two stages [33]: first, attention is distributed uniformly over the visual scene of interest; then, it is concentrated to a specific area (i.e., it is focused) for gaining information [20]. In a skill-oriented RLfD process, the instructor demonstrates skills mainly through their body joint poses. The above mechanism thus corresponds to that the human joints of interest are tracked uniformly at the initial stage, and then one joint providing the most information for learning is picked as an attention point. As for demonstration learning, the more predictable/track-able a body joint movement is, the less information the robot could gain from that part, and consequently, less attention the robot should pay to it. In other words, if a body joint moves out of expectation the most among all joints, it will be worth paying attention to.

To this end, we use the particle filter (PF) as it is robust and effective in predictions [37] and tracking [6]. In short, PF is a Bayesian filter which uses a group of samples to approximate the true distribution of a state [68]. Particularly, given the state observations, PF employs many samples (called *particles*) to describe the possible distribution of that state. The particles are denoted as

$$X_t := x_t^{[1]}, x_t^{[2]}, \dots, x_t^{[M]} \quad (1)$$

here M is the number of particles in the particle set X_t . Each particle $x_t^{[m]}$ (with $1 \leq m \leq M$) is a hypothesis as to what the true state might be at time t , and is first produced by a prediction model $p(x_t|z_{1:t-1})$ which is based on all history observations $z_{1:t-1}$, i.e., $x_t^{[m]} \sim p(x_t|z_{1:t-1})$. At each updating stage, particle $x_t^{[m]}$ is then re-sampled according to the *importance weight* $w_t^{[m]}$, i.e., the probability that the particle $x_t^{[m]}$ is consistent with the current observation z_t , i.e., $w_t^{[m]} = p(z_t|x_t^{[m]})$. In other words, each $x_t^{[m]}$ survives into the next stage with the probability $w_t^{[m]}$. For more details on the particle filter, refer to [68].

We apply one PF to track each relevant joint during the human demonstration. Specifically, state $\mathbf{x}_t^{[m]} \in \mathbb{R}^3$ describes the joint position in the sensor frame. We assume the state transits with additive Gaussian noise:

$$\mathbf{x}_t^{[m]} \sim \mathbf{x}_{t-1}^{[m]} + \Delta_{t-1} + \mathcal{N}(\mathbf{0}, \sigma_t \mathbf{I}) \quad (2)$$

where $\mathbf{x}_t^{[m]}$ denotes the predicted joint position vector, Δ_{t-1} is the observed joint shift: $\Delta_{t-1} = J_{t-1} - J_{t-2}$ (J_t refers to the observed joint position at time t); and $\mathcal{N}(\mathbf{0}, \sigma_t \mathbf{I})$ is the multivariate normal distribution with zero mean and diagonal covariance matrix $\sigma_t \mathbf{I}$. The *importance factor* for each particle is defined to be exponential to the Euclidean distance

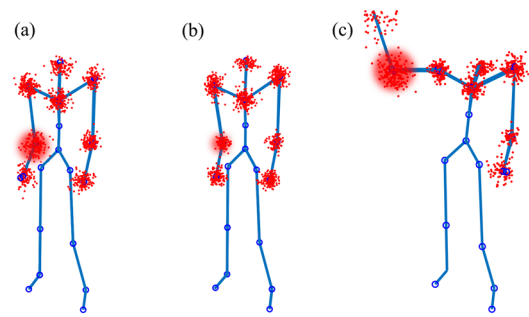


Fig. 2 The particle clouds evolve: **a** all clouds are initialized at the same size; **b** if the joint movement is small, the cloud shrinks: the picked cloud becomes smaller since the elbow did not move; **c** if the joint moves out of its cloud region, the cloud grows to catch the movement: the picked cloud becomes larger to adapt to the elbow's movements

between the predicted and observed joint position:

$$w_t^{[m]} = \eta e^{-2(\mathbf{x}_t^{[m]} - J_t)^T (\mathbf{x}_t^{[m]} - J_t)} \quad (3)$$

where η is the normalizer. Each joint in the body pose is tracked by a *particle cloud*, a group of particles \mathbf{X}_t . In order to dynamically adjust the cloud size in accordance with the joint movement, the variance σ_t is set to be proportional to the average Euclidean distance between the predicted and observed joint position:

$$\sigma_t = \frac{\alpha}{M} \sum_m [(\mathbf{x}_t^{[m]} - J_t)^T (\mathbf{x}_t^{[m]} - J_t)] \quad (4)$$

where α is a hyper-parameter and M is the number of particles. The σ_t indicates the cloud size: the greater the σ_t is, the more attention the robot should pay to the associated joint. Thus, the joint with maximum σ_t corresponds to the attention point. In the experiment, the α is set to 0.02 for the best tracking of human joints.

Figure 2 illustrates how the PF works to generate attention. The particle cloud functions as the robot's prediction of the joint future movements and is subject to change based on the current observations. Initially, the robot predicts the movements of all body joints of interest to be the same, i.e., all clouds are of the same sizes. During a demonstration, when a joint moves out of its cloud region, beyond the robot's prediction, the cloud grows to catch that movement and the robot will thereafter be likely to pay attention to that joint. Likewise, if the joint movement is small, within the robot's prediction, or no movement at all, the cloud shrinks, and chances are small that the attention will be given to that joint. Overall, the cloud size indicates the predictability of the instructor's body movements as well as the level of attention the robot needs to pay. At each time, the joint with the biggest cloud is picked as the attention point. This process loops with every new body pose as shown in Fig. 3.

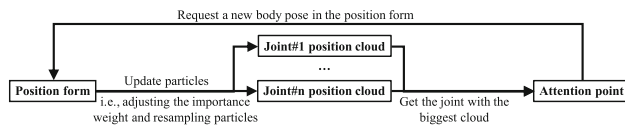


Fig. 3 The flow chart of the *Instant attention*

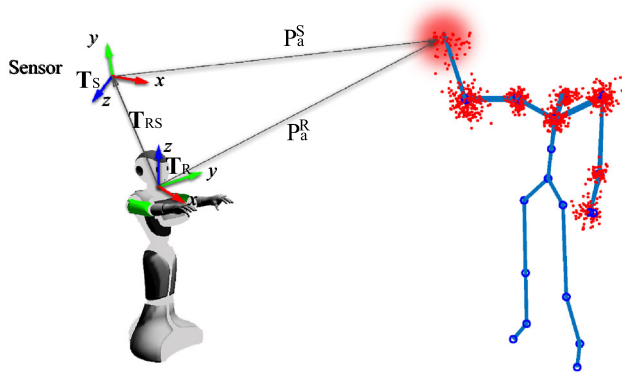


Fig. 4 The attention point P_a^S is located in the sensor frame T_S . We need to do the transformation $P_a^R = T_{RS}P_a^S$ to get P_a^R in the robot head frame T_R , where T_{RS} is the transformation from T_S to T_R

We now present a practical algorithm for *Instant attention* to generate gaze engagement instantly for robots (Algorithm 1). The algorithm takes *TrackedJoints* $JSet_{tracked}$ and the *BodyPose* in the position form $[J_t^{(1)}, J_t^{(2)}, \dots, J_t^{(n)}]$ as input, and outputs one attention point at each time. Specifically, the *TrackedJoints* contains the joints required to be tracked. In practice, the joints to be tracked are task-dependent, and should be defined according to the possible attention points on the instructor's body. For example, a cooking robot may only need to track the instructor's upper body movements and the joint correspondence can be configured by the developers based on the robot's physical structures. Another input *BodyPose* is the human body pose in the position form. The algorithm runs as follows: first, it initializes a particle filter with the same covariance for each tracked joint (line 2–4). Then it estimates the distribution of the next joint position (line 9–11), followed by the estimation correction given the current position observations (line 12–13). Finally, the algorithm adjusts the covariance of the noise distribution to capture the joint movement (line 14), and the attention point is found by selecting the joint with the maximum covariance value (line 15).

Once an attention point is generated, say P_a , it is worth mentioning that P_a is located in the sensor frame. To obtain the accurate attention point of the robot, a further transformation is required. Figure 4 illustrates how to transform P_a in the sensor frame T_S into the robot head frame T_R given the transformation T_{RS} from T_S to T_R .

The *Instant attention* method has several advantages. First, unlike other mechanisms (saliency-based, object-based, or

Algorithm 1: Instant attention

Input: TrackedJointSet $JSet_{tracked}$; BodyPose $[J_t^{(1)}, J_t^{(2)}, \dots, J_t^{(n)}]$, where $J_t^{(i)}$ is the 3D position of i -th joint at time t

Output: AttentionPoint $P_a \in \mathbb{R}^3$

```

[1] begin
[2]   for each joint  $j$  in  $JSet_{tracked}$  do
[3]     initialize  $j$ -th particle filter for joint  $j$ ;
[4]     initialize  $\sigma_t^{(j)} = 1$  for joint  $j$ ;
[5]   for each joint  $J_t^{(i)}$  in  $[J_t^{(1)}, J_t^{(2)}, \dots, J_t^{(n)}]$  do
[6]     if  $J_t^{(i)}$  is in  $JSet_{tracked}$  then
[7]        $\Delta_{t-1}^{(i)} = J_t^{(i)} - J_{t-1}^{(i)}$ ;
[8]       obtain particles  $\mathbf{x}_{t-1}^{[m]}$  from  $i$ -th particle filter;
[9]       for  $m = 1$  to  $M$  do
[10]        sample  $\mathbf{x}_t^{[m]}$  based on
[11]         $\mathbf{x}_t^{[m]} = \mathbf{x}_{t-1}^{[m]} + \Delta_{t-1}^{(i)} + \mathcal{N}(\mathbf{0}, \sigma_t \mathbf{I})$ ;
[12]        /* Equation2 */
[13]        calculate  $w_t^{[m]} = \eta e^{-2(\mathbf{x}_t^{[m]} - J_t)^T (\mathbf{x}_t^{[m]} - J_t)}$ ;
[14]        /* Equation3 */
[15]        for  $m = 1$  to  $M$  do
[16]          Select particles  $\mathbf{x}_t^{[m]}$  with probability  $\propto w_t^{[m]}$ ;
[17]        update
[18]         $\sigma_t^{(i)} = \alpha / M * \sum_m [(\mathbf{x}_t^{[m]} - J_t)^T (\mathbf{x}_t^{[m]} - J_t)]$ ;
[19]        /* Equation4 */
[20]    $P_a = \operatorname{argmax}_{J_t^{(i)}} \sigma_t^{(i)}$ ;
[21] return  $P_a$ ;

```

event-based), this method utilizes the particle cloud to track the instructor's joint movements, and automatically produces attention points based on the information gained from the movements. Second, the attention point is generated and shifted with very little abruptness because the spatial size of the cloud evolves smoothly. Specifically, the particle distribution $p(X_t)$ is iteratively sampled based on their previous distribution $p(X_{t-1})$ by the *importance weight* w_t , i.e., a $p(\mathbf{x}_{t-1}^{[m]})$ in X_{t-1} survives into X_t with probability $w_t^{[m]}$, even if the joint moves abruptly (i.e., $\mathbf{x}_t^{[i]} - J_t$ is large).⁵ Second, the joints to be tracked can be dynamically changed, offering a flexible and adjustable attention mechanism based on the RLfD task. Furthermore, if an object is involved in the demonstration and needs to be considered, we can simply interpret the object as an additional “joint”, or point of interest (POI) in a general term, for which a new set of cloud points is used to track its movements. The particle-based engagement

⁵ In practical implementation, a particle set may collapse to a single data point, which is called particle deprivation. Such problems only tend to arise when the number of particles is small relative to the state space. On the other hand, there are many ways to overcome this and one popular solution is to add a small number of randomly generated particles into the set after each resampling process, regardless of the actual observations and weights. See [68] section 4.3.4 for more detailed discussion.

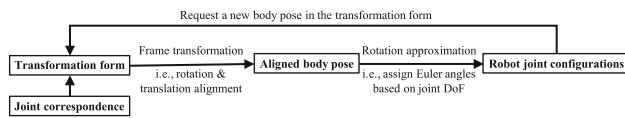


Fig. 5 The flow chart of the *Approximate imitation*

generation mechanism also applies to more general cases. For example, the joints considered in the evaluation experiments could potentially be generalized to a wide range of objects, each of which could be a POI and thus need to be tracked by a set of particle points.⁶

3.3 Approximate Imitation

Behavior imitation in robotics is usually formulated as an optimization problem, which needs to find the joint correspondence first [4], and then solves the inverse kinematics for the robot structure [24]. Both of the processes are difficult, computationally intensive, and robot-configuration-dependent, hence not applicable for generating imitation behavior for robots with different configurations and hardware. On the other hand, psychological results reported that people mimic behavior to communicate engagement by adopting similar postures or showing similar body configurations according to the context [14]. We thus relax the behavior imitation in robotics as follows: First, the robot is not required to search blindly for the best joint correspondence since the joint correspondence is task-dependent. We allow the user to explicitly specify the joint correspondence according to the RLfD context. Second, for those robot joints whose Degree of Freedom (DoF) do not match the human joint, we only set the joint angles for the available robot joints to approximate the human movements. Though this solution of approximation may not be optimal in the sense of behavior mimicry, it runs very fast (in real-time) to generate gesture engagement, achieving a balance between simplicity and optimality.

To achieve this, we propose the algorithm *Approximate imitation*, which allows robots to generate similar motions as the demonstrator for specified joints. Given the joint correspondence, the algorithm runs with two steps: frame transformation, and rotation approximation, as presented in Fig. 5.

The frame transformation is to transform the instructor's body pose to match the robot frames. To be specific, we leverage the transformation form of body poses to decompose the frame matching into two steps: first, rotation alignment and then translation alignment. The rotation alignment is to rotate the human joint frames so that their axes are aligned with the

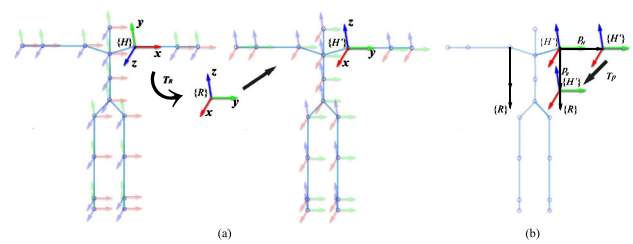


Fig. 6 Frame transformation. **a** Rotation alignment: aligning the local frame $\{H\}$ of the human body pose with the corresponding robot joint frame $\{R\}$ by rotation matrix R . The aligned local frame is $\{H'\}$. **b** Translation alignment: translating $\{H'\}$ in its parent frame by T_p to match the corresponding robot frame $\{R\}$ so that the human pose link p_H is aligned with the robot link p_R

robot joint frames, as shown in Fig. 6a; the translation alignment is to translate the human joint frames in their parent frames so that the initial skeletal structure of the demonstrator's body matches the robot initial configurations, as shown in Fig. 6b. To sum up, we represent the rotation alignment as T_R in the joint frame, $\{H\}$, and the translation alignment as T_p in the parent frame of $\{H\}$, $\{H_p\}$ (both represented in Homogeneous transformation). Then for $\{H\}$, its frame transformation is $T_H^{H_p} T_p \{H\} T_R$, where $T_H^{H_p}$ is the transformation from $\{H_p\}$ to $\{H\}$.

Since the DoF of the robot joint may not equal the DoF of its corresponding human joint, we could not have the exact movement mapping. Instead, we use the robot joint to approximate the human joint rotations as follows. First, a human joint rotation is converted into Euler forms, $(\theta_{roll}, \theta_{pitch}, \theta_{yaw})$. Second, if the DoF of a robot joint is 3 (roll, pitch and yaw) and exactly matches the human DoF, then the conversion is straightforward: rotate for the robot joint with roll first, then pitch, and finally yaw. If the DoF of a robot joint is 2 (e.g., roll and pitch), then the conversion can be approximated as rotating with roll first, and then pitch. If the DoF of a robot joint is 1 (e.g., roll only), then rotate with roll only. For example, in Fig. 7, the robot arm has the same structure as the demonstrator's but with different joint DoF, as shown in Fig. 7a and b. It can approximate the instructor's left arm movement by first converting T_S (the rotation) into Euler angles $(\theta_{roll}, \theta_{pitch}, \theta_{yaw})$, and then setting the joint roll to θ_{roll} , and the joint pitch to θ_{pitch} for the shoulder, ignoring the θ_{yaw} , as shown in Fig. 7c.

We now present the algorithm *Approximate imitation* in Algorithm 2. The algorithm takes joint correspondence *JointCorrespondence*, and instructor's body pose *JointMovement* in transformation form as input, and outputs the joint configurations, *JointConfigs*, for the robot. Specifically, *JointCorrespondence* defines the joint mapping, $\{J_i^H \rightarrow J_i^R\}$, from human joint J_i^H to robot joint J_i^R for part joints. The *JointMovement* is represented as a series of transformations along the skeletal structure, $[T_1, T_2, \dots, T_n]$ (see Sect. 3.1 for more

⁶ Instant attention can be summarized as the robot should pay attention to where it is uncertain, and uncertainty is measured using a particle filter; similar ideas have been generally explored in active learning and human studies [29,57].

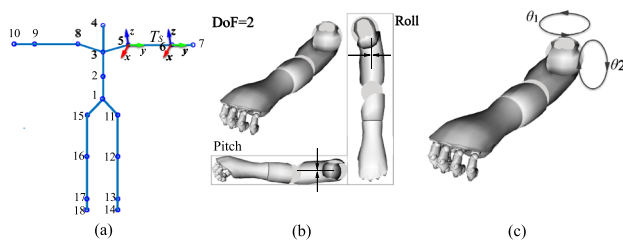


Fig. 7 Rotation approximation: **a** the instructor's left shoulder has a DoF of 3 and its transformation is T_S ; **b** the robot shoulder joint has a DoF of 2: *roll* and *pitch*; **c** the robot rotates for its shoulder the roll joint with θ_{roll} and then the pitch joint with θ_{pitch} , without considering θ_{yaw}

Algorithm 2: Approximate imitation

Input: JointCorrespondence
 $\{J_1^H \rightarrow J_1^R, J_2^H \rightarrow J_2^R, \dots, J_n^H \rightarrow J_n^R\}$,
 JointMovement= $[T_1, T_2, \dots, T_n]$

Output: JointConfigs q_R

```

[1] begin
[2]    $q_R = []$ ;  $Rotate\_align = []$ ;  $Translate\_align = []$ ;
[3]   /* following alignment only runs once */;
[4]   for ( $J_i^H \rightarrow J_i^R$ ) in
[5]      $\{J_1^H \rightarrow J_1^R, J_2^H \rightarrow J_2^R, \dots, J_n^H \rightarrow J_n^R\}$  do
[6]        $Rotate\_align.append(rotateAlign(J_i^R, J_i^H))$ ;
[7]        $Translate\_align.append(translateAlign(J_i^R, J_i^H))$ ;
[8]   for  $i$  in  $[1, 2, \dots, n]$  do
[9]      $T'_i = Translate\_align[i] * T_i * Rotate\_align[i]$ ;
[10]     $(\theta_{roll}, \theta_{pitch}, \theta_{yaw}) = convertToEuler(T'_i)$ ;
[11]    if  $DoF(J_i^R) == 3$  then
[12]       $append[\theta_{roll}, \theta_{pitch}, \theta_{yaw}]$  to  $q_R$ ;
[13]    else
[14]      if  $DoF(J_i^R) == 2$  then
[15]         $append[\theta_{roll}, \theta_{pitch}]$  to  $q_R$ ;
[16]      else
[17]        if  $DoF(J_i^R) == 1$  then
[18]           $append[\theta_{roll}]$  to  $q_R$ ;
[19]  return  $q_R$ ;

```

details). The algorithm runs as follows: first, it calculates the frame transformations from J^H to J^R , and saves the rotation alignment and translation alignment in *Rotation_align* and *Translation_align* (line 3–5). Then for each joint movement T_i in $[T_1, T_2, \dots, T_n]$, the algorithm transforms it into the corresponding robot frame T'_i by translation and rotation alignment, followed by a conversion into the Euler form (line 7–8). The algorithm proceeds by selecting the right angles from θ_{roll} , θ_{pitch} , and θ_{yaw} for the robot joint according to the DoF of the robot joint (line 9–16). The joint configurations are saved in q_R , and returned as the final output.

The *Approximate imitation* method has several advantages for generating imitation behavior for robots in RLfD. First, this algorithm runs in real-time as the imitation is only partially taken place on the instructor's body poses. In particular,

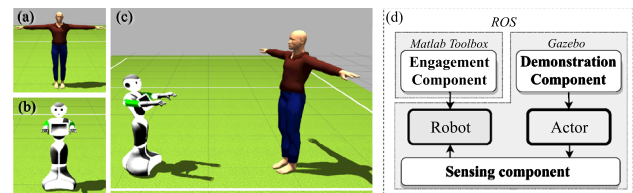


Fig. 8 RLfD simulation platform: **a** the simulated human instructor; **b** the virtual Pepper robot; **c** the instructor and robot are facing towards each other for teaching and learning; **d** platform composition

we take advantage of local transformations of body poses to avoid solving inverse kinematics for the whole robot joints, which is computationally intensive and may also not have closed-form solutions. Also, instead of finding the exact mapping for robot joint angles, we set configurations based on the DoF of the robot joint to achieve a similar motion trend. This conversion may sometimes distort movements, but, still, the directions and trends are captured (as reflected in 4). Second, this method is generic and applicable to standard skill-oriented RLfD. Depending on the RLfD scenario, we can also assign different joint correspondences to do a partial imitation. For other types of RLfD, e.g., object-related demonstrations or goal-oriented learning from demonstrations, we can also apply the proposed method to generate the approximate imitation based on the object or the goal. Specifically, we can replace the joint transformations with the poses of the object or the goal, and generate the target θ_{roll} , θ_{pitch} , and θ_{yaw} . Then we can adopt the inverse kinematic solvers to calculate a set of joint configurations to move the robot's end-site to the target pose $(\theta_{roll}, \theta_{pitch}, \theta_{yaw})$. Based on the DoF and the space constraints of the robot end-effectors, we can make similar approximations to have the end-effector only achieve the *roll* pose, the *roll* and *pitch* pose, or the complete target pose.

4 Evaluation

This section first introduces our RLfD simulation platform, then describes a preliminary study for determining the timing of imitating behavior, and finally presents the main user study.

4.1 RLfD Simulation Platform

Our RLfD simulation platform is composed of a virtual human instructor and a robot, as shown in Fig. 8a and b. The virtual human instructor performs different yet controlled types of movement skills, while the robot (a Pepper) needs to capture motion and learn skills from the instructor. Both parties stand facing each other in a simulated 3D space, as shown in Fig. 8c.

The simulation platform has three major components: demonstration component, sensing component, and engagement component, as shown in Fig. 8d.⁷ The demonstration component determines what movements the instructor needs to perform. We exploit motion capture (MoCap) data to simulate real movements. The MoCap data are recorded by 3D motion capturing systems with high precision and are usually used for simulations and animations [22]. The sensing component serves as a pose sensor, extracting body poses from the virtual instructor. This component also converts body poses between two representations (global positions and local transformations). Finally, the engagement component controls the robot's engagement communication. Based on the proposed algorithms, the robot could choose one of the three ways to communicating engagement in RLfD: showing attention (*A-mode*), showing imitation (*I-mode*), and showing both (*AI-mode*). We further add one more mode, i.e., no engagement (*N-mode*), to evaluate the effectiveness of these three modes. In *N-mode*, the robot just stands near the instructor and remains stationary without any body movements. Compared with the *A-mode*, the robot's gaze is fixed on the demonstrator's face and is not affected by the demonstrator's body movements. One might wonder why in *N-mode* the robot is not focusing randomly on one of the participant's joints. We argue that such setup is just a variant of gaze engagement except that the robot acts much less intelligently as it randomly moves its head, which could further deteriorate the human perception per se. In a human learning scenario, if a teacher sees students randomly move their heads and/or bodies, the teacher is likely to feel that the students may be listening but are quite lost—not paying attention to the right places—and thus gets quite confused about the students' actual learning status.

In this simulated RLfD, the tasks for robots to learn are sports skills performed by a virtual instructor. We chose sports skills for robots to learn as this type of movement has often been adopted in RLfD [9,31]. Four types of sports movements, i.e., boxing, rowing, swimming, and frisbeeing, are selected from CMU Graphics Lab Motion Capture Database⁸ as these four sports involve movements of various body parts. Regarding the policy deriving algorithms, even the state-of-the-art method may fail to deliver good learning outcomes, which may, in turn, change human participants' perception towards the demonstration gathering. Thus, to minimize any side-effects or biases introduced by the performance of the learning algorithms, we do not utilize any learning algorithms, and the robot has no actual learning abil-

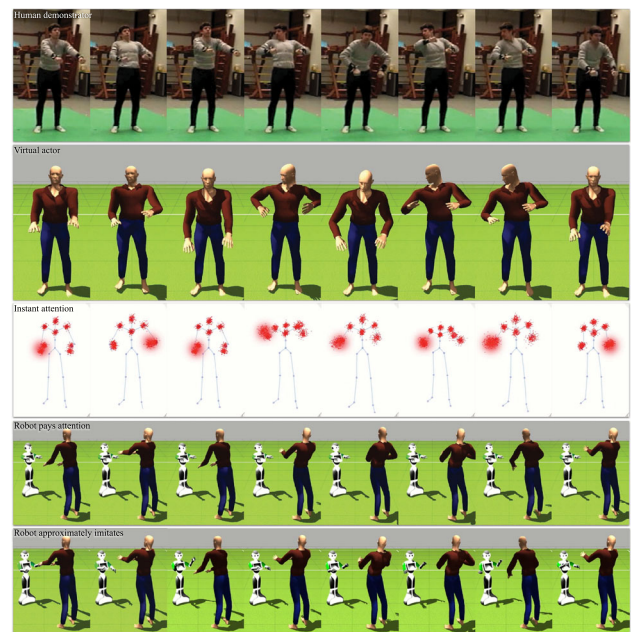


Fig. 9 An example to show how the platform works: Row 1 shows the human instructor's real demonstration; Row 2 shows the re-targeted demonstrations onto the virtual instructor; Row 3 and 4 present the running of *Instant attention* and robot showing attention (*A-mode*); and Row 5 presents the corresponding imitation engagement of the robot (*I-mode*)

ity in the demonstration gathering process. In the other words, the robot only communicates its engagement when observing the human demonstrations by showing different cues and will not learn the sports skills in the following experiments and studies.

Figure 9 presents an example of how the simulation platform works. The first row shows the human instructor's real demonstration, which is then re-targeted onto the instructor, as shown in the second row. The third and fourth rows present the running of *Instant attention* and robot showing attention (*A-mode*). The last row presents the approximate imitation behavior of the robot (*I-mode*). We purposely rotate the 3D scene in the last two rows to get a better view of the robot communicating engagement.

We chose an online simulation rather than a field test due to the following constraints and concerns: First, due to the current limitations of RLfD techniques, the demonstrators are usually required to wear motion-capture devices, confined in a designated space, and repeatedly showcase the target movements. This could potentially impact on their interaction with robots and their perception of the robot's behavior. Also, even if the current state-of-the-art vision-based methods, obtaining full-body motions with good precision (such that it could be used for task learning) is still very challenging. First, one camera Kinect is not enough to recover the body motion precisely. Second, multiple cameras also require cross-camera calibration, which itself could be hard to set up on the fly in

⁷ We checked the video sources and confirmed that the motion capture dataset can be copied, modified, or redistributed without permission. We also sent an email to the dataset correspondence and received their approval on using the original video in the paper.

⁸ <http://mocap.cs.cmu.edu/>.

an open-world HRI. Furthermore, as an initial attempt, we would like to study if humans could make sense of the learning engagement cues in a controlled environment without distractions and complications introduced by noises, jerks, etc. If we confirm the effectiveness of the proposed method in a controlled setting, we could then move to a realistic environment to test in the field. Also, it is very common these days to first train the algorithm on a simulation platform to reduce the costs. We thus use simulation in our experiment to avoid all these side effects and unexpected outcomes. Furthermore, we purposely select a viewpoint that allows the participants to have a better view of both the robot's and the instructor's behavior, i.e., the staging effect [67]. Second, the robot's engagement behavior could be evaluated in a more consistent and repeatable manner in a simulation. In a field test, the instructor's demonstrations are usually non-repeatable and could be easily influenced by robots' reactions. The simulation allows different engagement cues to be compared without bias. Second, the simulation provides a controllable and measurable environment to monitor and evaluate a system's performance from various metrics, which is often a necessity before algorithms are deployed in RLfD.

This simulation platform was built upon the Gazebo simulator⁹ and the Robot Operating System (ROS). We use the Matlab Robotics System Toolbox¹⁰ to facilitate the algorithm implementation.

4.2 Preliminary Study

In interpersonal communication, a person's imitation behavior, also called *mirroring behavior*, often happens after the partner's target behavior with a certain time delay [14,30]. In this paper, we generate such mirroring behavior via the approximation mechanism. We need to determine the exact time delay so that users can correctly recognize imitation as a learning engagement cue. We run a within-subject pilot experiment to check the appropriate timing of robot imitation relative to the target action.

Manipulated variable We set time delay as the independent variable in this study and experiment with three intervals: 0.5s, 1.0s, and 2.0s. Technically, we used a buffer to store instructors' body poses to postpone any imitation behavior. After proper setup, the buffer size was set to 15, 30, and 60 to achieve an appropriate time delay of about 0.5s, 1.0s, and 2.0s, respectively.

Subject allocation We recruited 30 participants (mean age: 35.5, female: 12) via Amazon Mechanical Turk (AMT) who had no prior experience with physical or virtual robots. Each participant watched three simulated RLfD videos cor-

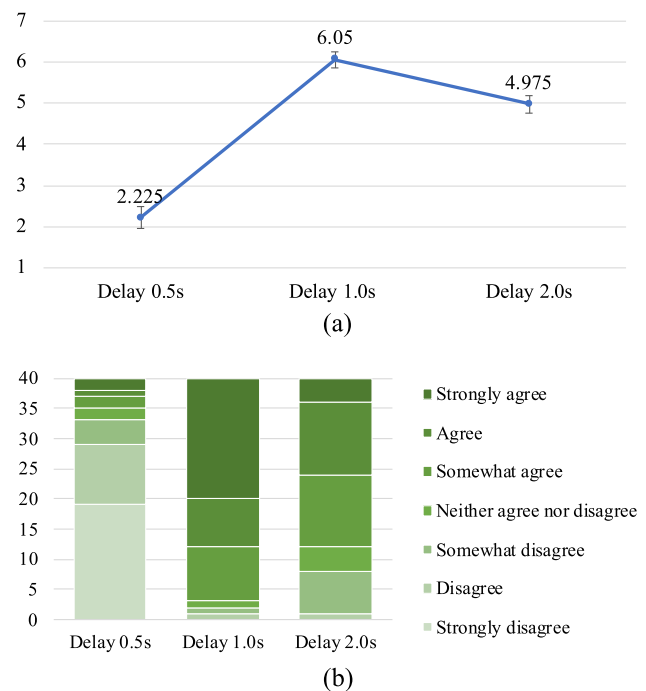


Fig. 10 Results for the right timing of behavioural engagement: **a** participants' ratings on robot learning behaviour, and **b** distribution of participants' feedback

responding to the three delay intervals. In the videos, the instructor was teaching the robot some type of sports skill, and we staged the 3D scene at a fixed angle for a better view of the robot imitations. We counterbalanced the presentation order of the different time delays. In other words, each subject tested only four out of the 16 possible combinations.

Dependent variables Participants watched videos showing the robot imitating the instructor with three different time delays. They were informed that the robot is supposed to learn sports skills from the demonstrator. After each video, they were asked to rate their agreement on a 7-point Likert scale as to whether the robot in the video is learning.

Figure 10 presents the average and overall rating distribution on different time delays. We run a repeated measures ANOVA with time delay as the factor, and find that there is a significant difference in delay-induced perception of robot learning engagement ($F(2, 58) = 88.37, p < 0.01, \eta^2 = .76$). Results of the Bonferroni posthoc test suggest that the engagement rating of delaying for 1.0s is significantly higher than that of delaying for 0.5s ($p < 0.01$) and 2.0s ($p < 0.01$). Overall, setting the imitation time delay to 1.0s can effectively communicate robots' learning engagement (70% agree and strongly agree). We apply this configuration to the *Approximate imitation* algorithm in the main user study.

One might be wondering that why the rating difference between 0.5s and 1s delay is noticeably dramatic, even larger than the difference between 1s and 2s delay. The cause

⁹ <http://gazebo.org/>.

¹⁰ <https://www.mathworks.com/hardware-support/robot-operating-system.html>.

may be the approximation mechanism adopted for generating the mirroring behavior. When the delay time is small (e.g., 0.5s), the approximate imitation algorithm generates the movement in a very responsive manner, almost at the same pace as the demonstrator's movement. The subjects are likely to feel that the robot is showing, rather than following, the demonstrator's movement. As the delay time becomes longer (e.g., 1s), the movement following effects becomes more obvious, and the robot appears to be learning from the demonstrator by mimicking his/her behavior. Consequently, the ratings between the 0.5s and 1s in terms of robot communicating learning engagement become higher. Such dramatic rating difference also confirms the necessity and importance of using the preliminary study to determine the appropriate delay time for the followed studies. Furthermore, different ratings of timing for robots to conduct mirroring behavior confirms that it is not the complexity of behavior that leads to different perceptions of robots' learning ability. Rather, it is when the behavior should be performed (i.e., engagement) that matters.

4.3 Main Study

To evaluate the effectiveness of engagement communication and our proposed cues on participants' perception of the robot and the demonstration, we conducted a within-subject experiment on an RLfD simulation platform, with an additional "without engagement" condition (*N-Mode*) as the baseline.

4.3.1 Hypothesis

Our proposed methods generate different types of engagement cues for robots to express their engagement. Accordingly, we first hypothesize that:

H1 (1) Regardless of actual cues taken, robots that communicate engagement are perceived to be significantly more engaged (*H1a*) in learning, and their learning behavior is significantly more socially acceptable (*H1b*) than those in the *N-mode*. Further, **(2)** imitation cue will receive a significantly higher engagement rating than attention cue (*H1c*), while combined cues will be rated significantly the most (*H1d*). Similarly, **(3)** imitation cue will be rated significantly more acceptable than attention cue (*H1e*) while combined cues will be rated significantly the most (*H1f*).

According to educational theory postulating that learners' engagement cues, especially gesture engagement, could have reciprocal effects on instructors [58], we hypothesize that:

H2 Robots communicating engagement via different cues will have significantly different influences on human participants. Specifically, **(1)** regardless of the cues, communicating engagement will significantly influence humans' estimation of the robot learning capability (*H2a*), and significantly raise the humans' expectations towards the learning outcomes

(*H2b*) than no communication. Further, **(2)** imitation cues will lead to a significantly higher estimation of the robot's capabilities than attention cues (*H2c*) while combined cues have the most significant influence than others. (*H2d*). Similarly, **(3)** imitation cues will result in a significantly higher expectation towards the learning outcome than attention cues (*H2e*) while combined cues have significantly the highest expectation than others (*H2f*).

We further hypothesize that the robot showing different engagement behavior can affect humans' assessment of demonstration quality. More specifically:

H3 (1) Regardless of the exact demonstrations shown to robots, different engagement cues will influence the human participants' assessment of the demonstration quality. Specifically, demonstrations for robots with attention cues, imitation cues, and hybrid cues will be rated as significantly more appropriate (in terms of the expected robot capabilities) than that without engagement cues even if they are the same (*H3a*). Further, **(2)** demonstrations for robots with imitation cues and the hybrid cues will have a significantly higher rating on appropriateness than that with attention cues (*H3b*).

In the study, these different aspects were measured via post-study questions with 7-point Likert scale answers, as shown in Figs. 11 and 12. We derived these questions in the user study based on the previous research on human–robot interactions and robot learning. Specifically, the questions to measure robot communicating engagement are adapted from the engagement studies [59,64]; the questions to measure participants' expectations towards the robot learning capability are derived based on the studies on human expectations and assessment of human–robot collaborations [39]. In addition to those engagement-related questions, we asked several factual questions, including "what sports skill is the virtual human demonstrating?" and "what skill is the robot learning?", and open questions, including "why do you have such estimate of the likelihood about the robot mastering the sports skill?" and "any comments on the robot's behavior". These factual and open questions were asked in each round to collect participants' understanding of the study materials and opinions on the robot learning. We also took two steps to ensure the effectiveness of the answers to all the questions. First, the questions could only be answered after participants took the necessary actions to understand the experiment. For example, the questions to measure engagement were only visible when the participants finished watching the full learning videos; and the questions to measure the participants' expectations also require the participants to provide the answers and their reasons (those without giving reasons could not proceed to the next questions). Second, all answers were manually checked to reject any invalid responses, e.g., a response with the same answers to all questions, and a response with vague and inconsistent comments. In total, we only removed 4 responses and the comments in those

responses: OK (subject#12), I do not care (subject#5), none of my business (subject#2) and I do not like robots (subject#23).

4.3.2 User Study Design

The study consisted of five sessions: one introductory session and four experimental sessions. The introductory session requested demographic information and presented a background story to engage users: the participant has a robot team of four for an Olympic game. They needed to assess the robots' performance when they were under a professional coach's tutelage. This session also presents all robots at the same time and showcase their learning capability by showing two video footages: one shows a different set of demonstrations given by a human demonstrator (dribbling a basketball) and another shows how the robot could learn such dribbling motion (learned by the behavioral cloning method). The participant can only proceed to the next sessions after finishing watching these two videos. This introductory session ensures that participants have been fully aware of the robot's learning capabilities. In experiment sessions, participants watched the human instructor's movements first and then monitored the robot learning process in the RLfD simulation platform. After each session, participants were required to fill post-study questionnaires. Each session checked one mode, and modes were counter-balanced with learning tasks. Specifically, we randomized the order of engagement modes and the four physical skills to ensure the mode applies evenly across different skills and the skill also occurs evenly across different modes. We recruited 48 participants from Amazon Mechanical Turk (AMT) (mean age: 30.9, female: 6, no prior experience with teaching robots, and no participation in the preliminary study. (Note that some participants reported that they had prior experience interacting with physical/virtual robots as they mentioned in answers to the open-ended questions. Our participants are representative in terms of their exposure to HRI and RLfD, and our findings could potentially be generalized to young adults with good digital literacy. We acknowledge that our participants lack diversity in age and education level. We are interested in exploring how different user populations may perceive and react to RLfD in the future.) Each subject receives two dollars as compensation for his/her contribution.

During the experiment, we asked the participants to rate if they perceived the robot was paying attention or imitating based on its behavior. This served as the manipulation check for validity, ensuring that our designs indeed convey the intended type of engagement. Note that what we want to assess is that, knowing that the robot can learn the target skill if taught right, do participants think that the demonstrations given are good enough for the purpose based on their observation of the robot's learning engagement. Without

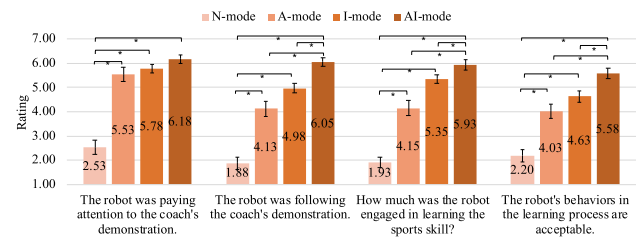


Fig. 11 Participants' ratings on robot engagement communications and their behavior in RLfD

emphasizing the robot behavior in the pilot study, participants rate the quality of demonstration based on their prior belief of whether or not they think the robot could acquire the demonstrated skill in an ideal situation. So we improved the question to specify the requirement that the quality of demonstrations should be assessed based on the robot's reactions (which is also highlighted in the experiments).

4.3.3 Analysis and Results

Manipulation check. The manipulation check for different engagement communications shows, in Fig. 11, that the manipulation is effective (for attention cue: repeated measures ANOVA, $F(3, 141) = 153.79$, $p < 0.01$, $\eta^2 = .80$; for imitation cue: repeated measures ANOVA, $F(3, 141) = 197.45$, $p < 0.01$, $\eta^2 = .84$). Robots in *A-mode* ($M = 5.53$, $SD = 1.85$) and *AI-mode* ($M = 6.17$, $SD = 1.11$) are indeed perceived to show more attention than robots in *N-mode* ($M = 2.53$, $SD = 1.83$); Bonferroni posthoc test $p < 0.05$. Also, more imitation behavior is reported by subjects with robots in *I-mode* ($M = 4.98$, $SD = 1.33$) and *AI-mode* ($M = 6.05$, $SD = 1.22$) than robots in *N-mode* ($M = 1.88$, $SD = 1.57$); Bonferroni posthoc test $p < 0.05$.

Efficacy of proposed engagement cues We analyze participants' ratings via a one-way repeated measures ANOVA with the mode as the independent variable. We find that both attention and imitation cues significantly improve the ratings of robots' engagement levels and their behavior, as shown in Fig. 11. Specifically, the robots with *A-mode* ($M = 5.53$, $SD = 1.85$), *I-mode* ($M = 5.78$, $SD = 1.03$) and *AI-mode* ($M = 6.17$, $SD = 1.11$) are perceived to be significantly more engaged in the learning process than the robot in *N-mode* ($M = 2.53$, $SD = 1.83$); repeated measures ANOVA, $F(3, 141) = 153.79$, $p < 0.01$, $\eta^2 = .80$. Thus, **H1a** accepted. Consequently, subjects accept the robots' behavior in RLfD (*A-mode*: $M = 4.02$, $SD = 1.78$, *I-mode*: $M = 4.62$, $SD = 1.44$, and *AI-mode*: $M = 5.58$, $SD = 1.39$) significantly more than the robot in *N-mode* ($M = 2.20$, $SD = 1.60$); repeated measures ANOVA, $F(3, 141) = 102.89$, $p < 0.01$, $\eta^2 = .73$. Thus, **H1b** accepted. Further, in terms of engagement, combined cues are reported to be significantly better than single cues; Bonfer-

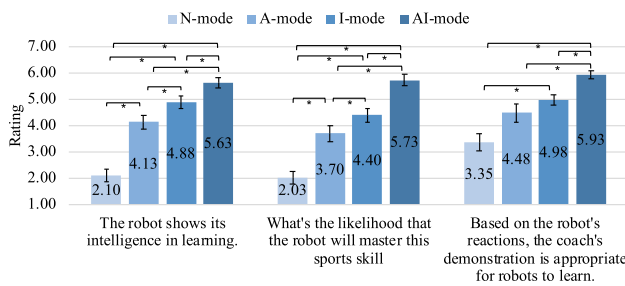


Fig. 12 Participants ratings on the effects of engagement communication on the participants' perception and their assessment of demonstration qualities

roni posthoc test $p < 0.01$. Thus, **H1d** accepted. in terms of acceptability, combined cues are reported to be significantly better than single cues; Bonferroni posthoc test $p < 0.01$. Thus, **H1f** accepted. However, we do not notice a significant difference between imitation cue and attention cue, thus **H1c** and **H1e** are both rejected. Therefore, **H1** is partially accepted.

Based on these analyses, we, therefore, conclude that:

Overall, our results partially support H1: showing attention, imitation, or both are perceived to be significantly more engaged in learning and is significantly more acceptable. Also, showing both behaviors is perceived to be significantly better than showing only one behavior. However, no significant difference can be found between showing attention and showing imitation.

Effects of engagement cues on participants' perception We then compare the effects of different engagement cues on subjects' perception via a one-way repeated measures ANOVA with the mode as the independent variable. In general, robot engagement communication significantly enhances the participants' estimation of robots' learning capabilities and the participants' expectation of the learning outcomes, even if none of the robots in the experiment have the learning ability (no learning algorithms are adopted in the user study). Specifically, in terms of estimating the robots learning capability, participants rated the robots, in Fig. 12, with respect to *A-mode* ($M = 4.13$, $SD = 1.70$), *I-mode* ($M = 4.88$, $SD = 1.49$) and *AI-mode* ($M = 5.63$, $SD = 1.21$) to be significantly more intelligent than the robots in *N-mode* ($M = 2.10$, $SD = 1.45$); repeated measures ANOVA, $F(3, 141) = 155.25$, $p < 0.01$, $\eta^2 = .80$. Thus, **H2a** accepted. Similarly, participants rated the robots with engagement behavior (*A-mode*: $M = 3.70$, $SD = 1.94$, *I-mode*: $M = 4.40$, $SD = 1.63$, and *AI-mode*: $M = 5.73$, $SD = 1.47$) to be more likely to master the skills than the robots without (*N-mode*: $M = 2.02$, $SD =$

1.59); repeated measures ANOVA, $F(3, 141) = 125.38$, $p < 0.01$, $\eta^2 = .76$. Thus, **H2b** accepted.

In addition, showing gesture engagement, i.e., *I-mode*, have significantly more influences on the participants than showing gaze engagement, i.e., *A-mode*. In particular, the robots in *I-mode* ($M = 4.88$, $SD = 1.49$) are perceived to be significantly more capable of learning the demonstrated skills than the robots in *A-mode* ($M = 4.13$, $SD = 1.70$); repeated measures ANOVA, $F(3, 141) = 155.25$, $p < 0.01$, $\eta^2 = .80$. Thus, **H2c** accepted. Similarly, the robots in *I-mode* ($M = 4.40$, $SD = 1.63$) receive significantly higher ratings than the robots in *A-mode* ($M = 3.70$, $SD = 1.94$) in terms of participants' expectation towards the learning outcomes; repeated measures ANOVA, $F(3, 141) = 125.38$, $p < 0.01$, $\eta^2 = .76$. Thus, **H2e** accepted.

Further, we also notice significant differences between robots in *AI-mode* and robots in other modes. Specifically, robots in *AI-mode* show significantly more intelligence in learning ($M = 5.63$, $SD = 1.21$) than robots in *N-mode* ($M = 2.10$, $SD = 1.45$), *A-mode* ($M = 4.13$, $SD = 1.70$), and *I-mode* ($M = 4.88$, $SD = 1.49$); repeated measures ANOVA, $F(3, 141) = 155.25$, $p < 0.01$, $\eta^2 = .80$. Thus, **H2d** accepted. Also, the robots in *AI-mode* ($M = 5.73$, $SD = 1.47$) are estimated by the participants to be significantly more likely to master the skill than the robots in modes (*N-mode*: $M = 2.02$, $SD = 1.59$, *A-mode*: $M = 3.70$, $SD = 1.94$ and *I-mode*: $M = 4.40$, $SD = 1.63$). Thus, **H2f** accepted. Note that in all different engagement modes and different skill settings, the robots are equipped with no learning algorithms and thus have no actual learning abilities.

In general, participants perceived the robot showing gaze engagement to be quite different from the one showing gesture engagement. Specifically, for the open-ended question "why do you have such estimate of the likelihood about the robot mastering the sports skill?", participants stated that, when showing gaze only, the robot seems to be "attentive throughout the demonstration so it knows what to do" (P16), "listening so there is a possibility of mastering this sport skill" (P18), and "learning by watching" (P13). These comments suggest that the participants perceive the robot exhibiting gaze engagement as an attentive listener in a mental state of learning. For the same question, when showing gesture engagement, the robot appears to our participants as "doing so well at following the instructor" (P4), and "making attempts and ideally would get better over time" (P18). These comments imply that the participants are judging the robots learning capability based on the robot's motivation to perform like the instructors. These differences between two types of perceptions also justify our motivation to separate gaze engagement and gesture engagement in this study.

Overall, our results support **H2**: communicating engagement significantly influence the humans' estimation of the robots' learning capabilities, and significantly changes their expectation towards the final learning outcomes, even though none of the robots have the learning abilities. Moreover, the gesture engagement in RLfD, i.e., imitation, presents significantly more influence on the participants than the gaze engagement. Furthermore, communicating engagement via two cues at the same time have significantly more effects on participants than communicating engagement via a single cue.

Effects on participants' assessment of demonstration qualities Finally, we analyze the participants' ratings on the appropriateness of instructors' demonstrations. As shown in Fig. 12, no significant difference can be found between *A-mode* ($M = 4.48, SD = 2.10$) and *N-mode* ($M = 3.35, SD = 2.08$); **H3a** rejected. However, compared with *A-mode*, only *AI-mode* ($M = 5.93, SD = 1.00$) significantly improves the participants' assessment of demonstration quality in RLfD, Bonferroni posthoc test $p < 0.01$. Thus, **H3b** partially accepted. Note that in different engagement modes, the skills to be learned are all generated by the same set of MoCap data. Thus all demonstrations are of the same quality.

Overall, our results partially support **H3**: communicating gesture engagement or combined engagement will significantly improve participants' assessment of demonstration qualities, while showing attention cannot, even though all the demonstrations are of the same quality.

Further, in the comments collected from the user study, we found that most participants explicitly stated that the robots without gesture engagement may fail in learning, and accordingly, they were more likely to adjust future demonstrations when the robots communicated no engagement or only gaze engagement.

5 Discussion

5.1 Engagement Communication for Robots in RLfD

The Choice of Engagement Cue Should Consider the Nature of the Learning Task

Our results show that robots' gesture engagement is preferable to gaze engagement in a physical skill-oriented RLfD, which can probably be explained by the correspondence between the practice of RLfD and the cone of learning [19]. Cone of learning, a.k.a. pyramid of learning or cone of experience, depicts the hierarchy of learning through involvement in real experiences [19]. It proposes that visual

receiving (just watching the demonstration) is a passive form of learning, and learners can only remember half of the knowledge passing through this channel two weeks later. In contrast, "doing the real thing" is a type of active learning that leads to deeper involvement and better learning outcomes [19].

In RLfD, the basic task for robots is to derive a policy from demonstrations and then reproduce the instructors' behavior [4]. On the one hand, a robot's imitation behavior resembles this "behavior reproducing" process; it is thus deemed actively engaged in the learning process. On the other hand, although showing gaze engagement implies that the robot is involved in the visual receiving of instruction, it is still considered as a passive way to learn. Consequently, instructors may conclude that a robot showing gesture engagement will have a deeper understanding and better mastery of the skill than that showing gaze engagement. Moreover, by analyzing the quality gap between a robot's imitation behavior and the demonstration (behavior to be reproduced), instructors may have a more accurate assessment of the robot's learning progress. In a word, to design effective engagement cues for robots in RLfD, we need to consider the nature of the learning task.

Engagement communication should reflect robot's actual capabilities

In our study, we do not equip the robot with any actual policy derivation algorithm since we want to avoid the perception bias caused by the algorithm selection. In other words, the robot has no learning ability. Still, many subjects are convinced that robots with engagement communication (attention, imitation, or both) would finally master the skill. They hold such a belief even if some tasks are technically very challenging for robots to learn because of the correspondence problem, e.g., swimming. These findings suggest that engagement communication can affect instructors' mental model of the robot's capability and progress. There can be a misalignment between instructors' expectations and the actual development as shown in our study. If instructors shape their teaching according to an inaccurate mental model, frustration may occur later in the RLfD process. Hence, it is critical to ensure that a robot's communication of engagement reflects its actual capabilities (policy development in the case of RLfD). One possible direction is that the robot engagement communication should be driven by the learning progress. For example, we can define the robot's actual capability as what it has learned. In this sense, the robot could gradually show its engagement, e.g., show more mimic behaviors, during the learning process.

5.2 Limitations

We are motivated by the phenomenon of human tutelage, in which the teachers build their mental model of students

by observing learning behavior, especially their engagement cues. The teachers, in this process, have different roles: observer (to monitor how students do) and demonstrator (to intervene in the learning process by providing demos). Similarly, a teacher in an RLfD process is also assigned these roles. They could observe the learning process and intervene at the point when they thought the demonstrations need to be changed, which is directly validated through their answers to questionnaires after each session. We would like to answer: (1) whether human instructors can understand the learning engagement cues simulated on robots, (2) how they make sense of these cues, and (3) how the instructors' interpretation of these cues may affect their perception and teaching. Whether and to what extent instructors' consequent change of teaching could improve robots' learning algorithm is to be investigated in future work. In our current evaluation, we cannot evaluate whether and how the robot behavior leads to improved demonstrations that in turn result in better learning. How the demonstrations in turn affect the learning process could be a promising future study.

In addition, this work has several limitations. First, in our study, engagement communication is decoupled from the robot's actual learning process. However, in the human or animal learning, such communication is usually associated with the learning process. For example, a student making good progress tends to show more gesture engagement [58]. We will investigate how to couple the learning process with engagement communication in the future. Second, in this paper, we only consider two types of learning engagement cues, i.e., attention and imitation. In practice, human learners may employ more diverse cues, e.g., spatially approaching, etc. Third, the proposed methods, *Instant attention* and *Approximate imitation* are both based on the human body poses. They may not apply to the learning tasks which do not necessarily involve the demonstrator's body movements, e.g., object manipulations. For those tasks, designing a good mechanism to communicate the robot engagement is still an open question. Fourth, in this work, we only consider skill-oriented RLfD in which the robot has to master a skill taught by instructors. Other types of RLfD, e.g., goal-oriented RLfD in which the robot learns how to achieve a goal from human examples, are inherently different in task settings. Though the proposed method may work, we still need to evaluate their effects in future work. Fifth, several studies show that there are gender effects on non-verbal communication [26]. Therefore, people in different genders may have different understandings of the robot engagement communication in our studies. We leave investigation of possible gender effects as future work.

And, lastly, we conduct the user study in an online simulation environment without a further offline and real-time RLfD test. Though the simulation is common practice to evaluate the idea in RLfD, the participants do not have any

control over the teaching process. How the participants might reshape future demonstration based on the robot's engagement feedback needs further investigation.

6 Conclusion

In this work, we propose two methods (*Instant attention* and *Approximate imitation*) to generate robots' learning engagement in RLfD. The *Instant attention* method automatically generates the point of attention and the *Approximate imitation* method produces robot imitation behavior. Based on the two methods, we investigate the effects of three types of engagement communication (showing attention, showing imitation, and showing both) via a within-subject user study. Results suggest that the proposed cues enable robots to be perceived to be significantly more engaged in the learning process and behave significantly more acceptably in RLfD than with no engagement communication. Also, these engagement cues significantly affect the human participants' estimation of robots' learning capabilities and the participants' expectation of the learning outcomes, even though all the robots have no actual learning abilities. In particular, imitation cue influences instructors' perceptions significantly more than attention cue, while the hybrid cues significantly outperform a single cue. We also find that showing gesture or combined engagement significantly improves instructors' assessments of demonstration qualities. This paper takes the first step to reveal the potential effects of communicating engagement on humans in RLfD.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Abbeel P, Ng AY (2004) Apprenticeship learning via inverse reinforcement learning. In: Proceedings of the twenty-first international conference on machine learning. ACM, p 1

2. Admoni H, Scassellati B (2015) Robot nonverbal communication as an AI problem (and solution). In: 2015 AAAI fall symposium series
3. Anzalone SM, Boucenna S, Ivaldi S, Chetouani M (2015) Evaluating the engagement with social robots. *Int J Soc Robot* 7(4):465–478
4. Argall BD, Chernova S, Veloso M, Browning B (2009) A survey of robot learning from demonstration. *Robot Auton Syst* 57(5):469–483
5. Argyle M, Cook M (1976) *Gaze and mutual gaze*. Cambridge University Press, Cambridge
6. Arulampalam MS, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans Signal Process* 50(2):174–188
7. Atkeson CG, Schaal S (1997) Robot learning from demonstration. In: *ICML*, vol 97, pp 12–20
8. Bailenson JN, Yee N (2005) Digital chameleons: automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychol Sci* 16(10):814–819
9. Bentivegna DC, Ude A, Atkeson CG, Cheng G (2002) Humanoid robot learning and game playing using PC-based vision. In: *IEEE/RSJ international conference on intelligent robots and systems*, vol 3. IEEE, pp 2449–2454
10. Bernieri FJ, Rosenthal R (1991) Interpersonal coordination: behavior matching and interactional synchrony. *Fundament Nonverb Behav. Studi Emot Soc Int*
11. Breazeal C, Brooks A, Gray J, Hoffman G, Kidd C, Lee H, Lieberman J, Lockerd A, Mulanda D (2004) Humanoid robots as cooperative partners for people. *Int J Humanoid Robots* 1(2):1–34
12. Calinon S, Billard A (2007) Active teaching in robot programming by demonstration. In: *RO-MAN 2007-The 16th IEEE international symposium on robot and human interactive communication*. IEEE, pp 702–707
13. Chao C, Cakmak M, Thomaz AL (2010) Transparent active learning for robots. In: 2010 5th ACM/IEEE international conference on human–robot interaction (HRI). IEEE, pp 317–324
14. Chartrand TL, Bargh JA (1999) The chameleon effect: the perception-behavior link and social interaction. *J Personal Soc Psychol* 76(6):893
15. Chartrand TL, Maddux WW, Lakin JL (2005) Beyond the perception-behavior link: the ubiquitous utility and motivational moderators of nonconscious mimicry. In: *The new unconscious*, pp 334–361
16. Chernova S, Thomaz AL (2014) Robot learning from human teachers. In: *Synthesis lectures on artificial intelligence and machine learning*, vol 8(3), pp 1–121
17. Chernova S, Veloso M (2009) Interactive policy learning through confidence-based autonomy. *J Artif Intell Res* 34:1–25
18. Crick C, Osentoski S, Jay G, Jenkins OC (2011) Human and robot perception in large-scale learning from demonstration. In: *Proceedings of the 6th international conference on Human–robot interaction*. ACM, pp 339–346
19. Dale E (1969) *Audiovisual methods in teaching*. Dryden Press. <https://books.google.co.uk/books?id=s6ocAAAAMAAJ>
20. Eriksen CW, Hoffman JE (1972) Temporal and spatial characteristics of selective encoding from visual displays. *Percept Psychophys* 12(2):201–204
21. Fuente LA, Ierardi H, Pilling M, Crook NT (2015) Influence of upper body pose mirroring in human–robot interaction. In: *International conference on social robotics*. Springer, pp 214–223
22. Gleicher M (1998) Retargetting motion to new characters. In: *Proceedings of the 25th annual conference on computer graphics and interactive techniques*. ACM, pp 33–42
23. Gonsior B, Sosnowski S, Mayer C, Blume J, Radig B, Wollherr D, Kühnlenz K (2011) Improving aspects of empathy subjective performance for HRI through mirroring emotions. In: *Proceedings of the IEEE international symposium on robot and human interactive communication, RO-MAN 2011, Atlanta, USA*
24. Grochow K, Martin SL, Hertzmann A, Popović Z (2004) Style-based inverse kinematics. *ACM Trans Graph TOG* 23:522–531
25. Guthrie JT, Cox KE (2001) Classroom conditions for motivation and engagement in reading. *Educ Psychol Rev* 13(3):283–302
26. Hall JA (1978) Gender effects in decoding nonverbal cues. *Psychol Bull* 85(4):845
27. Hatfield E, Cacioppo JT, Rapson RL (1993) Emotional contagion. *Curr Dir Psychol Sci* 2(3):96–100
28. Ho J, Ermon S (2016) Generative adversarial imitation learning. In: Lee DD, Sugiyama M, von Luxburg U, Guyon I, Garnett R (eds) *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5–10, 2016, Barcelona, Spain*, pp 4565–4573. <https://proceedings.neurips.cc/paper/2016hashcc7e2b878868cbac992d1fb743995d8fAbstract.html>
29. Hogarth L, Dickinson A, Austin A, Brown C, Duka T (2008) Attention and expectation in human predictive learning: the role of uncertainty. *Q J Exp Psychol* 61(11):1658–1668
30. Hove MJ, Risen JL (2009) It's all in the timing: interpersonal synchrony increases affiliation. *Soc Cognit* 27(6):949–960
31. Ijspeert AJ, Nakanishi J, Schaal S (2002) Movement imitation with nonlinear dynamical systems in humanoid robots. In: *IEEE international conference on robotics and automation. Proceedings of the ICRA'02*, vol 2. IEEE, pp 1398–1403
32. Jackson SL, Krajcik J, Soloway E (1998) The design of guided learner-adaptable scaffolding in interactive learning environments. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co, pp 187–194
33. Jonides J (1983) Further toward a model of the mind's eye's movement. *Bull Psychon Soc* 21(4):247–250
34. Kim S, Kim C, You B, Oh S (2009) Stable whole-body motion generation for humanoid robots to imitate human motions. In: *IEEE/RSJ international conference on intelligent robots and systems. IROS 2009*. IEEE, pp 2518–2524
35. Koenemann J, Burget F, Bennewitz M (2014) Real-time imitation of human whole-body motions by humanoids. In: *2014 IEEE international conference on robotics and automation (ICRA)*. IEEE, pp 2806–2812
36. Koenig N, Takayama L, Mataric M (2010) Communication and knowledge sharing in human–robot interaction and learning from demonstration. *Neural Netw* 23(8–9):1104–1112
37. Konidaris G, Kuindersma S, Grupen R, Barto A (2012) Robot learning from demonstration by constructing skill trees. *Int J Robot Res* 31(3):360–375
38. Kuno Y, Sadazuka K, Kawashima M, Yamazaki K, Yamazaki A, Kuzuoka H (2007) Museum guide robot based on sociological interaction analysis. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, pp 1191–1194
39. Kwon M, Huang SH, Dragan AD (2018) Expressing robot incapability. In: *Proceedings of the 2018 ACM/IEEE international conference on human–robot interaction, HRI'18*. ACM, New York, pp 87–95. <https://doi.org/10.1145/3171221.3171276>
40. Laskey M, Chuck C, Lee J, Mahler J, Krishnan S, Jamieson K, Dragan A, Goldberg K (2017) Comparing human-centric and robot-centric sampling for robot deep learning from demonstrations. In: *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, pp 358–365
41. Li J, Ju W, Nass C (2015) Observer perception of dominance and mirroring behavior in human–robot relationships. In: *Proceedings of the tenth annual ACM/IEEE international conference on human–robot interaction*. ACM, pp 133–140
42. Lockerd A, Breazeal C (2004) Tutelage and socially guided robot learning. In: *Proceedings 2004 IEEE/RSJ international conference*

- on intelligent robots and systems, 2004. (IROS 2004), vol 4. IEEE, pp 3475–3480
43. Mutlu B, Forlizzi J, Hodgins J (2006) A storytelling robot: modeling and evaluation of human-like gaze behavior. In: 2006 6th IEEE-RAS international conference on humanoid robots. Citeseer, pp 518–523
 44. Mutlu B, Shiwa T, Kanda T, Ishiguro H, Hagita N (2009) Footing in human–robot conversations: how robots might shape participant roles using gaze cues. In: Proceedings of the 4th ACM/IEEE international conference on human robot interaction. ACM, pp 61–68
 45. Nagai Y, Muhl C, Rohlfing KJ (2008) Toward designing a robot that learns actions from parental demonstrations. In: IEEE International conference on robotics and automation, 2008. ICRA 2008. IEEE, pp 3545–3550
 46. Ng AY, Russell SJ et al (2000) Algorithms for inverse reinforcement learning. In: *Icml*, pp 663–670
 47. Pekrun R, Linnenbrink-Garcia L (2012) Academic emotions and student engagement. In: *Handbook of research on student engagement*. Springer, pp 259–282
 48. Pitsch K, Vollmer AL, Mühlhig M (2013) Robot feedback shapes the tutor’s presentation: how a robot’s online gaze strategies lead to micro-adaptation of the human’s conduct. *Interact Stud* 14(2):268–296
 49. Riek LD, Paul PC, Robinson P (2010) When my robot smiles at me: enabling human–robot rapport via real-time head gesture mimicry. *J Multimodal User Interfaces* 3(1–2):99–108
 50. Ross S, Gordon G, Bagnell D (2011) A reduction of imitation learning and structured prediction to no-regret online learning. In: *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. pp 627–635 (2011)
 51. Saunders J, Nehaniv CL, Dautenhahn K (2006) Teaching robots by moulding behavior and scaffolding the environment. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human–robot interaction*. ACM, pp 118–125
 52. Schaal S (1997) Learning from demonstration. In: Mozer M, Jordan MI, Petsche T (eds) *Advances in Neural Information Processing Systems 9*, NIPS, Denver, CO, USA, December 2–5, 1996, pp 1040–1046. <http://papers.nips.cc/paper/1224-learning-from-demonstration>
 53. Sena A, Zhao Y, Howard MJ (2018) Teaching human teachers to teach robot learners. In: 2018 IEEE international conference on robotics and automation (ICRA). IEEE, pp 1–7
 54. Sidner CL, Kidd CD, Lee C, Lesh N (2004) Where to look: a study of human-robot engagement. In: *Proceedings of the 9th international conference on intelligent user interfaces*. ACM, pp 78–84
 55. Sidner CL, Lee C, Kidd CD, Lesh N, Rich C (2005) Explorations in engagement for humans and robots. *Artif Intell* 166(1–2):140–164
 56. Silpasuwanchai C, Ma X, Shigemasa H, Ren X (2016) Developing a comprehensive engagement framework of gamification for reflective learning. In: *Proceedings of the 2016 ACM conference on designing interactive systems*. ACM, pp 459–472
 57. Silver D, Bagnell JA, Stentz A (2012) Active learning from demonstration for robust autonomous navigation. In: 2012 IEEE international conference on robotics and automation. IEEE, pp 200–207
 58. Skinner EA, Belmont MJ (1993) Motivation in the classroom: reciprocal effects of teacher behavior and student engagement across the school year. *J Educ Psychol* 85(4):571
 59. Strait M, Vujovic L, Floerke V, Scheutz M, Urry H (2015) Too much humanness for human–robot interaction: exposure to highly humanlike robots elicits aversive responding in observers. In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems, CHI’15*. ACM, New York, pp 3593–3602. <https://doi.org/10.1145/2702123.2702415>
 60. Suleiman W, Yoshida E, Kanehiro F, Laumond JP, Monin, A (2008) On human motion imitation by humanoid robot. In: *IEEE international conference on robotics and automation, 2008. ICRA 2008*. IEEE, pp 2697–2704
 61. Sun M, Ma X (2019) Adversarial imitation learning from incomplete demonstrations. *arXiv preprint arXiv:1905.12310*
 62. Sun M, Mou Y, Xie H, Xia M, Wong M, Ma X (2019) Estimating emotional intensity from body poses for human-robot interaction. *arXiv preprint arXiv:1904.09435*
 63. Sun M, Zhao Z, Ma X (2017) Sensing and handling engagement dynamics in human–robot interaction involving peripheral computing devices. In: *Proceedings of the 2017 CHI conference on human factors in computing systems*. ACM, pp 556–567
 64. Sun M, Zhao Z, Ma X (2017) Sensing and handling engagement dynamics in human–robot interaction involving peripheral computing devices. In: *Proceedings of the 2017 CHI conference on human factors in computing systems, CHI’17*. ACM, New York, pp 556–567. <https://doi.org/10.1145/3025453.3025469>
 65. Szafrir D, Mutlu B (2012) Pay attention!: designing adaptive agents that monitor and improve user engagement. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, pp 11–20
 66. Takayama L, Dooley D, Ju W (2011) Expressing thought: improving robot readability with animation principles. In: 2011 6th ACM/IEEE international conference on human–robot interaction (HRI). IEEE, pp 69–76
 67. Thomas F, Johnston O, Thomas F (1995) *The illusion of life: disney animation*. Hyperion, New York
 68. Thrun S, Burgard W, Fox D (2005) *Probabilistic robotics*. MIT Press, Cambridge
 69. Yamazaki K, Yamazaki A, Okada M, Kuno Y, Kobayashi Y, Hoshi Y, Pitsch K, Luff P, vom Lehn D, Heath C (2009) Revealing gauquinn: engaging visitors in robot guide’s explanation in an art museum. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, pp 1437–1446

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Mingfei Sun is a research associate at Oxford University (jointly with Microsoft Research Cambridge). He received the PhD degree in computer science and engineering from the Hong Kong University of Science and Technology in 2020. He is interested in Reinforcement Learning and Human-Robot Interaction.

Zhenhui Peng is a postdoctoral researcher from the Department of Communications and Network, Aalto University. He received his PhD degree in March 2021 from the Hong Kong University of Science and Technology. His research background is Human-Computing Interaction. He is particularly interested in understanding and facilitating communications on social media and developing intelligent agents to support thinking tasks including but not limited decision-making, problem-solving, critical thinking, and creativity.

Meng Xia is a postdoctoral researcher from the School of Computing, Korea Advanced Institute of Science and Technology. She received her PhD degree in Aug 2020 from the Hong Kong University of Science and Technology. Her research background is Data Visualization, Human-AI Interaction, and Education. Her research focuses on using visual analytics and emerging AI techniques (e.g., NLP, AR/VR) to

improve students' learning and instructors' teaching in an online educational environment.

Xiaojuan Ma is an associate professor of Human-Computer Interaction (HCI) at the Department of Computer Science and Engineering (CSE), Hong Kong University of Science and Technology (HKUST). She received the Ph.D. degree in Computer Science at Princeton University. She was a post-doctoral researcher at the Human-Computer

Interaction Institute (HCII) of Carnegie Mellon University (CMU), and before that a research fellow in the National University of Singapore (NUS) in the Information Systems department. Before joining HKUST, she was a researcher of Human-Computer Interaction at Noah's Ark Lab, Huawei Tech. Investment Co., Ltd. in Hong Kong. Her background is in Human-Computer Interaction. She is particularly interested in data-driven human-engaged AI (HEAI) and Human-Engaged Computing (HEC) in domains including but not limited to education, health, and design.