



Who to Observe and Imitate in Humans and Robots: The Importance of Motivational Factors

Johann Chevalère^{1,4}  · Murat Kirtay² · Verena V. Hafner^{3,4} · Rebecca Lazarides^{1,4}

Accepted: 31 August 2022 / Published online: 23 September 2022
© The Author(s) 2022

Abstract

Imitation is a vital skill that humans leverage in various situations. Humans achieve imitation by observing others with apparent ease. Yet, in reality, it is computationally expensive to model on artificial agents (e.g., social robots) to acquire new skills by imitating an expert agent. Although learning through imitation has been extensively addressed in the robotic literature, most studies focus on answering the following questions: *what* to imitate and *how* to imitate. In this conceptual paper, we focus on one of the overlooked questions of imitation through observation: *who* to imitate. We present possible answers to the *who*-to-imitate question by exploring motivational factors documented in psychological research and their possible implementation in robotics. To this end, we focus on two critical instances of the *who*-to-imitate question that guide agents to prioritize one demonstrator over another: outcome expectancies, viewed as the anticipated learning gains, and efficacy expectations, viewed as the anticipated costs of performing actions, respectively.

Keywords Social robotics · Educational psychology · Observational learning · Imitation · Motivation

1 Introduction

Rapid developments in social robotics and the accompanying integration of robots into human daily life yield important questions for the future of society. For instance, in the educational context, one may ask whether robots could one day assist teachers in the classroom by first learning from the teachers and then, by transferring the acquired knowledge and skills to students, as humans do. To answer this question, understanding how humans and robots learn from other agents [1] and especially by learning through imitation—an exceptionally developed capacity in humans [2, 3]—is of pri-

mary importance. We envision that social robots could rely on, for example, multiple human teachers and tutors and use the principles of observational learning (e.g., to observe and prioritize the demonstrators that would maximize learning gains) to extract the features that characterize human teaching and tutoring expertise.

Two families of imitation techniques can be identified in robotics, which have been developed respectively in the last two decades. First, robots can learn through means of demonstrations [4–8]. This approach, also known as programming through demonstration or imitation learning, employs action examples provided by a teacher or demonstrator, which reduces the complexity of search spaces for learning, thus overcoming the abortive slowness of learning through experience alone using standard learning methods, including reinforcement learning and its variations. However, the main downside of this approach is the requirement of a demonstrator’s kinaesthetic guidance, which places constraints on the learning situation (e.g., physical presence, availability of the demonstrator, a limited number of demonstrators), precluding the use of valuable resources such as videos [9–11].

A second, more recent family of imitation-learning techniques in robotics, drawing closer to the flexible and common ability of humans to rely solely on observations [12–16] is imitation learning from observation [10, 11, 17, 19]. In this

Johann Chevalère and Murat Kirtay contributed equally to this work.

✉ Johann Chevalère
johann.chevalere@hotmail.fr

- ¹ Department of Education, University of Potsdam, Potsdam, Germany
- ² Department of Cognitive Science and Artificial Intelligence, Tilburg School of Humanities and Digital Sciences (TSHD), Tilburg University, Tilburg, The Netherlands
- ³ Adaptive Systems Group, Department of Computer Science, Humboldt-Universität Zu Berlin, Berlin, Germany
- ⁴ Cluster Science of Intelligence (SCIoI), Technische Universität Berlin, Berlin, Germany

perspective, the agent learns to perform imitation learning via perceptual information only, without action guidance, where the robot learns (through inverse reinforcement learning [10, 17]) or mirrors (based on the retrieval of prior visuo-motor associations [20–22]) the action policy from the perceptually extracted features. Because learning through observation offers the possibility to exploit a vast amount of virtual resources in addition to in-presence observations and social interactions, new challenges come into consideration regarding how to filter relevant information. In this regard, *who-to-imitate* becomes a central question, as selective social mechanisms may be needed to choose, among many exemplars, which demonstrations best align with the objectives of the learning activity, for example, through recognizing that specific demonstrators are not worth imitating. We note that the *who-to-imitate* question has received limited attention relative to the other fundamental questions of *what* and *how* to imitate, which are extensively discussed in the work of [8, 18, 23–26] and therefore, are not addressed here.

In this paper, we emphasize selective social learning as a supplementary avenue to circumvent the challenges in current imitation learning methods, envisioning that robots could be endowed with the strategic *choice* to follow a given demonstrator in line with imitation learning objectives. The rest of the manuscript is organized as follows: First, we briefly explore the concept of selective social learning in humans by discussing the attributes known to influence humans' preferences for certain informants and demonstrators. Then, we explore similar aspects of selective social learning in robotics. Finally, with the perspective of imitation learning from observation in mind, we propose a conceptual bridge between humans and robots on how to foster selective social imitation by exploring motivational factors, as possible answers to the *who-to-imitate* question.

2 Selective Social Learning

2.1 Selective Social Learning in Humans

Humans are selective social learners. A number of factors underlying human's ability to trust a message and/or learn from specific partners-while not from others-have been identified in the literature [27–30], for reviews). According to Mills [27], attributes pertaining to three domains underlay the ability to adopt a critical stance in children and adults: attributes relate to the informant, the message, and the target (e.g., the learner). First, as a general tendency, a message's acceptability by a target is more likely when an informant shows a high degree of expertise, which may be inferred based on age, authority, and familiarity, in addition to showing positive attitudes during the delivery of the message. Second, to be accepted as valid, the message itself should be

highly accurate, matching the informant's domain of expertise, and be generalizable to other domains. Third, the target's skills are needed to decode the informant's intentions, detect flaws, use prior knowledge, and be motivated to critically evaluate the reliability of the message.

Another classification is provided by Harris et al. [28], based on studies in children. The authors first describe a class of social-affective attributes of the informant-target interaction, where a message acceptability depends on the informant's effective use-and the learner's perception-of congruent affective signals, and an appropriate use of gestures and vocalizations. Second, cognitive attributes are proposed, including the learner's ability to reason about the informant, that is, to perceive the reliability and accuracy of the informant. Third, social-positional attributes include considerations of the informant's social standing and personality: positive attitudes, same in-group membership, familiarity, and high social status. Fourth, attributes concern the learner's ability to use prior knowledge in the evaluation of the message. For instance, if the message contradicts the learner's strong intuitions or expectations, it is less likely to be accepted as valid.

Focusing on situations involving infant learners, Poulin-Dubois and Brosseau-Liard [29] propose a distinction in terms of cues that directly or indirectly evidence the competency of the informant. First, the direct cues include the degree of accuracy of informants during the learning task, where infants attend to and learn more (e.g., in a word-learning task) with informants that were previously accurate, who exhibit a competent behaviour in the task (e.g., using an object in appropriate ways) and make a congruent use of emotional expressions (e.g., reacting positively to positive experiences). Second, the authors report cues that tend to evidence the competence of the informant more indirectly: when the information receives agreement by a third party, when the informant is older than the learner, or shows confidence rather than uncertainty.

Consistent with these studies, two broad classes of attributes can be delineated, echoing Koenig and Sabbagh's [30] distinction in terms of epistemic and non-epistemic attributes. First, epistemic attributes in the process of information transfer relate to the (a) degree of expertise of the informant, through evaluation of accuracy or knowledge level shown in (related) past or current tasks by the learner, and (b) the learner's ability to use prior knowledge. Second, non-epistemic attributes pertain to (c) the social-positional domain (e.g., group membership, age, social standing) and (d) social-affective processes (e.g., decoding emotional cues, inferring confidence).

2.2 Selective Social Learning in the Context of Imitation Learning from Observation

Regarding imitation more specifically, a growing body of studies on motor learning has documented the attributes that drive the learner's preference towards certain demonstrators with the aim of reproducing/imitating a given behaviour through observation ([31, 32] for reviews). Note that the focus made here on observational learning does not undermine the existence nor the importance of other forms of social learning in humans, such as instructed and collaborative learning [33, 34].

In line with the literature on message acceptability, the observation and selection of specific demonstrators by humans in the field of motor learning is motivated by factors that largely relate to epistemic considerations including the expertise level of the demonstrator [31, 32, 35, 36], in addition to learners' self-assessment of task difficulty and competence level [37, 38]. Compelling evidence shows the role of a skilled demonstrator in enhancing learners' motor execution and retention of the to-be-learned behaviour. The benefit of a skilled demonstrator is explained in terms of a correct performance providing more accurate representations of the task [31, 32, 36]. Despite this general trend, the benefit of an unskilled demonstrator compared to a skilled one is sometimes reported, in particular, in that it promotes cognitive engagement and error detection [31, 32, 36, 39]. Similarly, a so-called coping model, transiting from being unskilled to being skilled on a given task, may enhance the observers' affective/cognitive outcomes such as self-efficacy and perceived task difficulty [31, 32, 39, 40]. For example, human learners holding doubts about their ability to learn a given task may be more interested in peer models (who share similar characteristics in terms of age and perceived competence) and may learn more by observing them gradually progress in a task than by observing an expert demonstrator directly [39–41]. Another finding concerns the benefits of observing a combination of demonstrators with different skill levels on skill execution and retention compared to a single demonstrator [31, 32, 36].

Taken together, studies on selective social learning crucially stress epistemic factors as a major driving force behind a learner's preference towards specific informants and demonstrators. As we can see in the next section, robotic studies similarly place great emphasis on epistemic considerations to guide the selection of learning partners, that is, those partners who would maximize the robots' internal reward and learning gains [42–44].

2.3 Selective Social Learning in Robots

Selective social learning is an important concept that allows the robots to differentiate reliable and unreliable information

sources (e.g., interaction partners) and leverage information provided by the reliable sources to perform a given task. Here we focus on selective learning studies in the context of forming varying degrees of trust in interaction partners and selecting the trustworthy partner based on a robot's internal signals: increasing the cumulative reward, decreasing computational cognitive load, to mention a few. Trust studies in the literature have been addressed in three distinct research directions: human trust in robots, robot trust in humans, and reciprocal trust. Most human trust-in-robots studies focus on determining core components of trust, including predictability of the robot, anthropomorphism of the robot design, etc. [45, 46]. Although human trust in robots is the dominant research direction in the literature, there have been notable attempts to propose robot trust models by using epistemic attributes of the robot and the partner and reciprocal trust between humans and robots [47, 48].

Here, in line with the present paper's scope, we review the studies that address robot trust in interaction partners, including simulated agents, humans, and robots, by using the epistemic attributes. Kirtay et al. [49] presented a robot trust model on the Nao robot in an interactive sequential pattern recalling tasks with online pre-programmed partners that have reliable, unreliable, and random strategies to guide the robot. In this setting, interaction partners guide the Nao robot to perform visual recall tasks [47]. For instance, a reliable partner will provide useful suggestions to the robot that enables the robot to process the visual patterns that were associated with a low cost (i.e., low computational load incurred on the robot to recall a visual pattern). In turn, interacting with reliable partners yield increments in cumulative reward during the experiment. After conducting the experiments with all partners, the authors provide a free-choice to the robot to select the trustworthy interaction partner who reduces the robot's cognitive load in performing the interactive task. We note that the authors also extended the same robot trust model in multimodal human–robot and robot–robot interaction settings [50, 51]. Chen et al. [52] presented a computational model that infers human trust by a robot in a collaborative tabletop task (i.e., cleaning a table) between a robot arm and a human partner. Overall, the study concludes that the robot with the trust model leads to a better task performance by using the cumulative reward relative to the robot without the trust model. Patacchiola and Cangelosi [53] proposed a robot trust model implemented on the iCub robot. In a first object naming learning phase, the robot builds prior knowledge by receiving accurate information linking objects to correct labels. Then, in a familiarisation phase, reliable and unreliable informants respectively provide correct and incorrect labels. Lastly, in an explicit judgement phase, the robot assesses the trustworthiness of the partners and recognizes the reliable partner by using its prior knowledge acquired during the learning phase.

Overall, robot trust studies are dedicated to provide computational accounts of a robot's trust, based on the interaction partners' epistemic attributes during learning tasks (e.g., accuracy and reliability of the information by the partner and the robot's performance) that influence the social interaction dynamics (e.g., increasing the cumulative reward).

2.4 Selective Social Learning in the Context of Imitation Learning from Demonstration

Despite a growing body of research devoted to imitation learning from observation ([10, 11, 18, 19, 54], there is still, to our knowledge, a lack of research investigating selective social learning in this context. There are, however, some few studies using the principle of selective social learning in the context of learning from demonstration. For example, a study by Nguyen and Oudeyer [44] combined intrinsic motivation and socially guided learning on a robot arm to explore and learn different motor movements (e.g., placing a fishing can on a task space area) to complete a fishing task. In this setting, the robot either performed autonomous exploration or imitated teachers' policies via kinaesthetic teaching to reach a specific position on the surface area. The active strategic choice of the robot to either self-explore or rely on social guidance at each learning episode was based on the previously developed intrinsic-motivation algorithm [55], that is, an architecture guiding the robotic agent to prioritize options decreasing sensorimotor prediction error in reaching a given outcome and, thus, maximize learning progress.

More importantly, the authors also emphasized the *who-to-imitate* question by using the previous intrinsic-motivation architecture to integrate demonstrations from multiple teachers in order to determine which teacher is the most expert and should be imitated. Three demonstrators provided kinaesthetic information with variable levels of expertise on two types of tasks (i.e., throwing or placing a ball with a fishing rod). Applying the principle of prediction error minimization to differentiate between the three demonstrators for each type of task, the SGIM-ACTS architecture learned who to ask for demonstrations among the available teachers by deciding which demonstrator yielded the biggest learning progress for each task when imitated.

In the next section, we build further on the idea of social selective learning by focusing on motivational factors documented in psychological research to give insights on how to further develop current approaches to learning from observation in robotics. Owing to the ubiquitous role of epistemic attributes in guiding human and robots' preferences for specific demonstrators and informants, in the following, we will focus on motivational factors that take these considerations

into account.

 Springer

3 Motivational Factors Guiding Observational Learning

3.1 Expectancies Related to Outcomes and Self-efficacy

In humans, an important part of the motivation process at play when learning from observation is the estimation of activity-related expectations in terms of positive outcomes and self-efficacy, according to the Social Learning Theory [56, 57]. While outcome expectancy refers to an individual's evaluation that the behaviour will lead to a certain outcome (i.e., *Would imitating this demonstration effectively lead me to hand write the letter "a"?*), efficacy expectation is the belief that one can successfully perform the behavior required to produce the outcome (i.e., *Will I be capable of handwriting the letter "a" considering the difficulty to imitate this demonstration?*). It is assumed that individuals develop and regulate belief in their ability to ensure that acts or events occur, which acts as an expected reinforcement through increased internal reward (i.e., a positive internal reinforcement).

However, the positive reward associated with successful observational learning may come at a cost related to the characteristics of the task in terms of difficulty and time, according to the Expectancy-Value Theory [58, 59]. Pursuing a costly activity is thought to depend on how much effort can be deployed considering task difficulty, which follows a non-monotonic relationship [60]. This means that as long as the objective remains achievable from the agent's perspective, higher perceived expectancies (in terms of actual outcomes and self-efficacy) lead to greater internal motivation, allocation of cognitive resources, intensity, and persistence at the service of the task [60–62]. However, beyond acceptable levels of task difficulty, the relationship between effort and task difficulty may turn into a negative reinforcement for the agent as the efforts are no longer efficient, leading to task disengagement [60].

Therefore, the notion of cost is particularly essential for determining task choice, as it reflects the amount of "sacrifice" that an agent can tolerate when pursuing the desired learning activity. It has strong implications for selective social imitation learning in humans and robots, as it allows to determine whether a demonstrator is worth being imitated and if so, the duration and strength of an agent's motivation in engaging in that imitation.

3.2 Benefit-over-cost

To determine whether a particular demonstrator is worth being imitated, a decision-making process might need to take place resulting from a benefit-over-cost analysis, understood as weighing the anticipated costs of effort against the anticipated benefits [62–64]. One important element to consider

in the benefit-over-cost analysis is the amount of irrelevant and/or missing information in the demonstration. Cognitive psychology has long documented that suppressing irrelevant information consumes cognitive resources [65–68] and a recent study investigated this question with regard to the benefit-over-cost analysis in human learning. Using computational models based on data from a reinforcement learning task embedded with task-irrelevant information, Sidarus et al. [69] revealed that decision making was affected and learning reduced as resulting from conflict costs related to irrelevant information processing being traded off against expected rewards. From a robotic perspective, the notion of cost can be easily understood in terms of computational resources expenditure during task completion. For example, facing an unclear or incomplete demonstration, the observer robot might need to allocate a great amount of resources to clarify and/or fill incomplete parts of the demonstration [54] which could be avoided by favoring another demonstrator. Therefore, the anticipation of imitation-related costs may represent a crucial component of a robot's observational learning and, a fortiori, be useful for the purpose of selective social imitation learning.

4 Towards Connecting Who to Imitate with Motivation

Here we suggest that learning from observation in robotics may benefit from leveraging motivational factors [70] with the idea that an agent is most motivated when the task to be learned is both achievable and yields high learning gains [55, 71]. The conceptual contribution of this paper proposes that imitation-related outcome expectancy, in terms of anticipated learning gains, be weighed against self-efficacy expectancy, in terms of the anticipated costs required to overcome the demonstrator's unavailable/infeasible actions. This benefit-over-cost analysis could be applied sequentially to several demonstrators, whose outputs (i.e., ratio) would characterize their respective worthiness for imitation and thus the motivation for the robot to prioritize specific demonstrators. We emphasize that this could apply to both virtual and physically-present demonstrators if the robot is equipped with cameras to record the demonstrator's movements, thus relying on visual information throughout the imitation process. Kinaesthetic teaching is viewed as an efficient resource that could complement this approach. Nevertheless, we note that our aim is to provide insights to help address the *who*-to-imitate question, not to propose an approach for *how* to imitate. In the next section, we will use the example of learning how to hand-write the letter "a" to illustrate our purpose, as this entails direct applications for the educational context where social robots might be particularly useful for tutoring physical skills [72].

4.1 Outcome Expectancy Viewed as Anticipated Learning Gains

A simple way to consider artificial outcome expectancy is as resulting from the comparison between three visual patterns (i.e., images), which output value would represent the anticipated learning gains in case the imitation of the demonstrator is to be performed. Using our example of hand-writing the letter "a", a first image may correspond to the robot's prior knowledge about the goal (cf. the upper right square in panel a, Fig. 1), in other words, how a prototypical letter "a" should look like in the robot's memory.

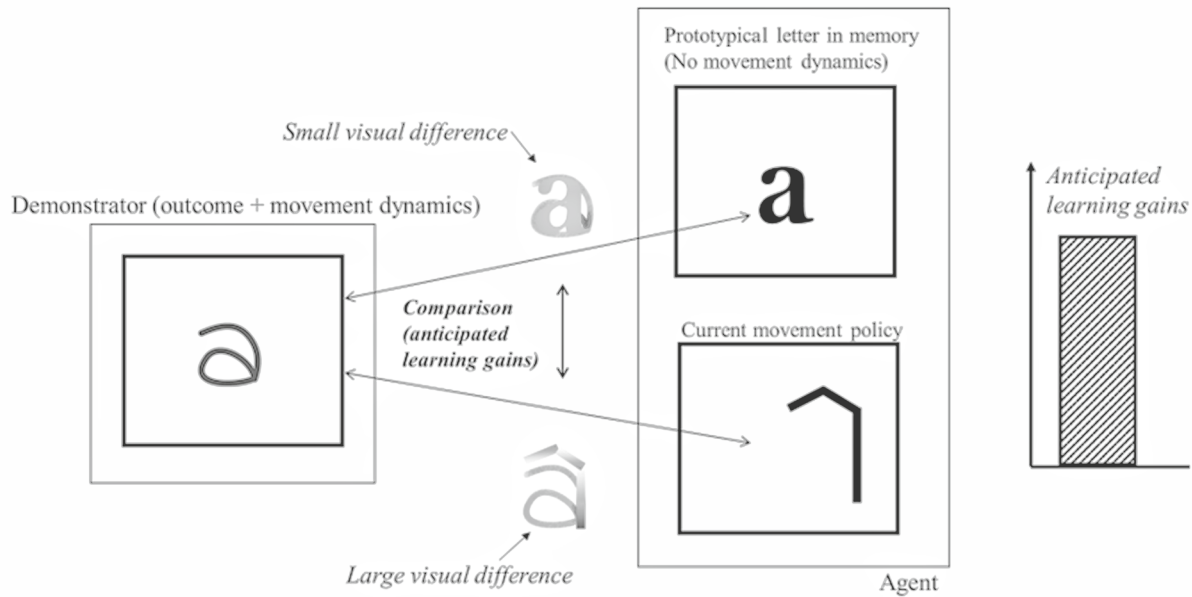
The second image may correspond to the robot's level of competence, that is, the final letter that the robot actually generates with its own policies (e.g., learned through prior goal-babbling activities) (cf. the bottom right square in panel a, Fig. 1). The third image may correspond to the demonstrator's final letter, that is, how the letter "a" looks at the end of the demonstration (cf. the left square in panel a, Fig. 1). The latter image could be retrieved from the recorded demonstration using an event detection system searching for the closest representation of the letter "a" contained in the robot's memory. The anticipated learning gains could be a function that accepts the visual pattern generated by the demonstrator, the robot's prototypical letter, and the pattern drawn by the robot. This function compares these three patterns to assess whether imitating the demonstrator's movement trajectories would potentially lead to minimizing the competence-related difference (i.e., the anticipated learning gain value) between the robot's actual outcome and desired outcome, that is, better approximating the prototypical letter image held in memory.

4.2 Self-efficacy Expectancy Viewed as Anticipated Cost

One flexible aspect of observational learning is to actively determine whether the demonstration is actually reproducible by the observer. To this end, we envision that a self-efficacy expectancy function could serve this purpose. Self-efficacy expectancy is understood as the estimation of the anticipated costs during a first attempt to execute the imitation. Contrary to the anticipated learning gain values obtained from fixed images, the value for anticipated costs involves the temporal discourse of the demonstrators' actions. This would allow the robot to segment the demonstrator's movements into sub-goals (e.g., a given number of time epochs or segments) in order to distinguish feasible segments from infeasible ones in the demonstration [54, 73]. Figure 1, panel b., displays an example of a demonstration incompleteness related to invisible movement trajectories deployed by the demonstrator while hand-writing the letter "a".

After segmenting the demonstration into sub-goals, a total error (i.e., the anticipated costs value) calculated based on

a.



b.

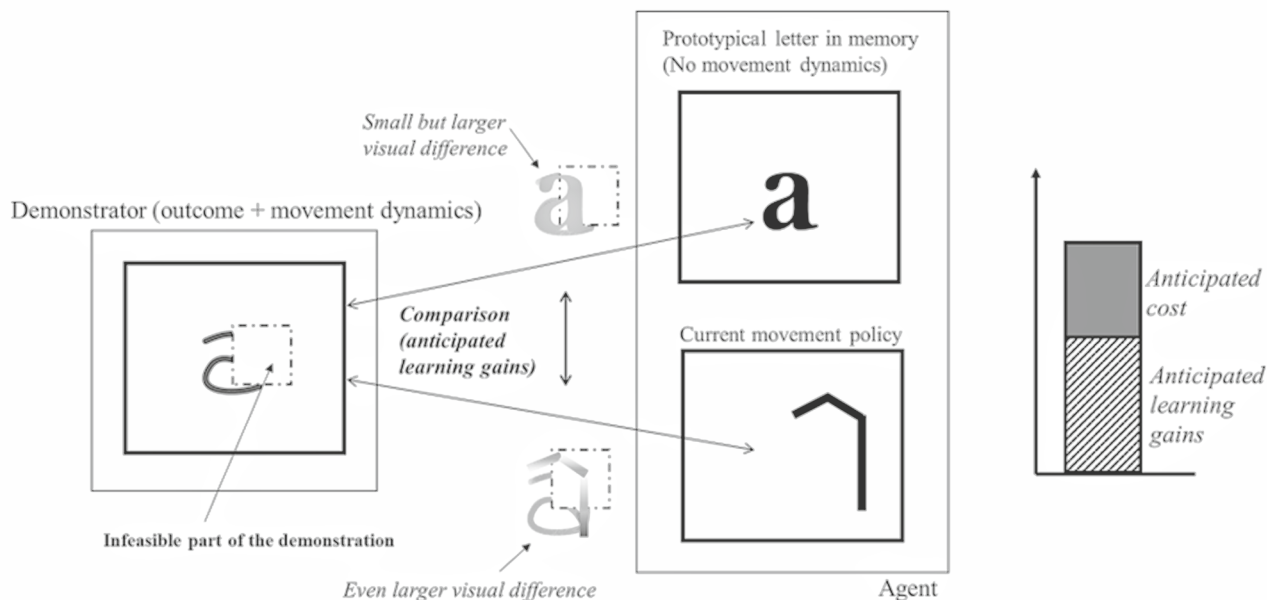


Fig. 1 Illustration of an Agent's Outcome Expectancy (learning gains: upper panel **a**) and Self-efficacy Expectancy (cost due to the incompleteness of the demonstration: lower panel **b**). *Note.* In panel **b**, the

imitation requires to fill in unfeasible parts of the demonstration, which consumes additional computational resources (i.e., anticipated costs) thus balancing out the anticipated learning gains displayed in panel **a**

the proportion of infeasible segments out of the total number of segments could be derived. For example, only 30% of the segments for a demonstrator A might be infeasible after execution attempts of all segment, compared to a prohibitive 80% for a demonstrator B. The anticipated cost value is then to be outweighed against the value for anticipated learning

gains in the benefit-over-cost analysis,¹ which will ultimately guide the selection and prioritization of demonstrators.

¹ To determine a benefit-over-cost ratio, the anticipated learning gains and cost values should be normalized on the same scale (e.g., 0 to 1). For example, a value close to one for anticipated cost could be illustrated by all parts of the demonstration being unfeasible while a value close to one for anticipated learning gains could be illustrated by a demonstrator's final letter closely resembling the letter held in the robot's memory.

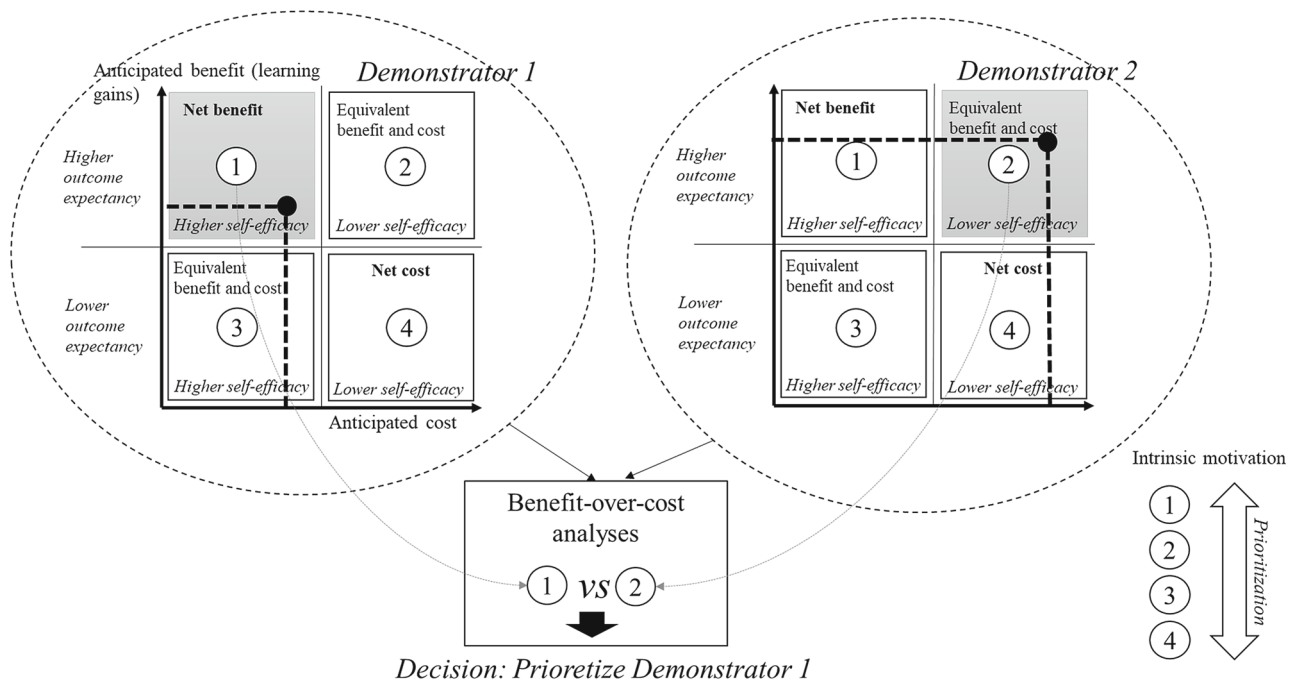


Fig. 2 Schematic representation of the selection process of a demonstrator during observational learning. *Note.* Even though Demonstrator 2 yields more promising learning gains (black dotted horizontal line), Demonstrator 1 is prioritized as it leads to a net benefit. If the first two prioritized demonstrators fall in the same category, a more fine-grained analysis may be conducted within this category. For the sake of simplicity, we include discrete categories (1 to 4, ranging from high to

low intrinsic motivation) to evaluate the worthiness of each demonstrator. Categories 2 and 3 are ranked by giving priority to the expected outcome (achieve high learning gains at the expense of many efforts) over expected self-efficacy (achieve small learning gains at the expense of little effort) but the priority could be inverted in case of time and resource constraints are put on the robot

4.3 Using the Benefit-over-cost Analysis to Guide the Selection of Specific Demonstrators for Imitation

The benefit-over-cost analysis, the proposed instance guiding the selection of specific demonstrators, corresponds to the ratio of anticipated learning gains and cost for each demonstrator separately, and is illustrated in Fig. 2.

For each demonstrator, the analysis may lead to four different scenarios. In scenario 1, important anticipated gains clearly outweigh the minimal cost required to imitate the demonstration, in other words, high outcome expectancies and high efficacy expectations, i.e., the most interesting option. In scenarios 2 and 3, there is no net benefit in pursuing the imitation. Scenario 2 corresponds to high outcome expectancies but low efficacy expectations (important learning gains are anticipated but the imitation is deemed difficult to follow). Conversely, Scenario 3 corresponds to low outcome expectancy but high self-efficacy (small learning gains are anticipated but the imitation is deemed easy to follow). In scenario 4, both outcome expectancies and efficacy expectations are low as the large anticipated costs clearly outweigh the small anticipated benefits, i.e., the least interesting option.

After this process is conducted for each demonstrator separately, a comparison of the benefit-over-cost outputs could be performed across the different demonstrators to determine which one should be selected and prioritized. The prioritized demonstrations may then be subjected to multiple imitation iterations by the robot to refine the movement trajectory.

5 Conclusion and Future Research Directions

This paper has elaborated on imitation learning and focused especially on *who-to-imitate*, a fundamental question for social learning yet overlooked in social robotics. We have placed a particular emphasis on learning from observation, where imitation does not necessarily depend on action guidance, which is viewed as a highly flexible and common human-like ability [12–16] that should be considered in social robots.

Considering possible applications for education, we have discussed the notions of outcome expectancy and self-efficacy expectancy as potent factors deployed during observation to orient agents to prioritize specific demonstrators for imitation. Although we introduce the possible connection of

observational learning with motivational factors for the purpose of hand-drawing letters, the same concepts could also be extended to different teaching purposes involving motor tasks: drawing specific shapes for the objects that have multiple parts (e.g., legs, hand, head for humans) and forming complex objects, (e.g., plane, by attaching lego-like small elements together), to mention a few.

Most of the attention of the present paper has been paid to epistemic considerations (i.e., the expertise level of the demonstrator and feasibility of the demonstration by the learner) to explain an agent's social selection, which are consistently put forward both in psychology and robotics studies. However, we acknowledge that non-epistemic considerations may also be highly relevant for imitation learning (i.e., in particular, the social-affective classes of attributes highlighted in Sect. 2). In this regard, imitation learning architectures could integrate some of these non-epistemic attributes as possible future research directions, for instance, via vicarious reinforcement to guide the selection of demonstrators [74, 75]. This may correspond to the use of affective feedback provided on the demonstrator's performance to infer the consequences that pursuing the imitation would entail for the robot. For example, a demonstrator's failures suggested by the expression of negative emotions might be used as a warning to either discard the demonstrator or, more interestingly, to identify specific errors that should be avoided [76] which is especially shown in human coping models [31, 32, 36]. That direction is particularly promising considering the advances made on the recognition of emotions both on faces and body poses in robotics [77, 78], as well as on the modelling of artificial empathy [79–81], which are yet to be connected with imitation learning.

Acknowledgements None.

Author Contributions Initiative: RL and VH; Funding acquisition: RL and VH; Conceptualization: JC; Literature search: JC and MK; Manuscript writing: JC and MK; Critical revision: MK, JC, RL and VH. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2002/1 “Science of Intelligence”—project number 390523135.

Data Availability No datasets were generated or analyzed during the current study.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethical approval This study is part of a research project that received an approval from the University of Potsdam Ethics Committee (number 82/2020).

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

- Kirtay M, Chevalère J, Lazarides R, Hafner VV (2021) Learning in social interaction: perspectives from psychology and robotics. In 2021 IEEE international conference on development and learning (ICDL), pp 1–8. Doi: <https://doi.org/10.1109/ICDL49984.2021.9515648>
- Farmer H, Ciaunica A, Hamilton AFDC (2018) The functions of imitative behaviour in humans. *Mind Lang* 33(4):378–396. <https://doi.org/10.1111/mila.12189>
- Dautenhahn K, Nehaniv CL (2022) Imitation in animals and artifacts. *Boston Rev.* <https://doi.org/10.7551/mitpress/3676.001.0001>
- Argall BD, Chernova S, Veloso M, Browning B (2009) A survey of robot learning from demonstration. *Robot Auton Syst* 57(5):469–483. <https://doi.org/10.1016/j.robot.2008.10.024>
- Atkeson CG, Schaal S (1997) Robot learning from demonstration. In *ICML*, vol 97, pp 12–20. <http://www-clmc.usc.edu/publications/A/atkeson-ICML1997.pdf>
- Billard A, Calinon S, Dillmann R, Schaal S (2008) *Survey: robot programming by demonstration*. Springer, Berlin, pp 1371–1394. https://doi.org/10.1007/978-3-540-30301-5_60
- De Rengervé A, Hirel J, Andry P, Quoy M, Gaussier P (2011) On-line learning and planning in a pick-and-place task demonstrated through body manipulation. In 2011 IEEE international conference on development and learning (ICDL). IEEE, vol 2, pp 1–6. Doi: <https://doi.org/10.1109/DEVLRN.2011.6037336>
- Ravichandar H, Polydoros AS, Chernova S, Billard A (2020) Recent advances in robot learning from demonstration. *Annu Rev Control Robotics Auton Syst* 3:297–330. <https://doi.org/10.1146/annurev-control-100819-063206>
- Zhou L, Xu C, Corso JJ (2018) Towards automatic learning of procedures from web instructional videos. arXiv preprint [arXiv:1703.09788](https://arxiv.org/abs/1703.09788). <https://doi.org/10.48550/arXiv.1703.09788>
- Torabi F, Warnell G, Stone P (2018) Generative adversarial imitation from observation. *ArXiv180706158*. <https://doi.org/10.48550/arXiv.1807.06158>
- Torabi F, Warnell G, Stone P (2019) Recent advances in imitation learning from observation. *ArXiv190513566*. <https://doi.org/10.48550/arXiv.1905.13566>
- Kline MA (2015) How to learn about teaching: an evolutionary framework for the study of teaching behavior in humans and other animals. *Behav Brain Sci* 38:E31. <https://doi.org/10.1017/S0140525X14000090>

13. Meltzoff AN (1988) Imitation of televised models by infants. *Child Dev* 59(5):1221. <https://doi.org/10.1111/j.1467-8624.1988.tb01491.x>
14. Meltzoff AN (1988) Imitation, objects, tools, and the rudiments of language in human ontogeny. *Hum Evol* 3(1):45–64. <https://doi.org/10.1007/BF02436590>
15. Marshall PJ, Meltzoff AN (2014) Neural mirroring mechanisms and imitation in human infants. *Philos Trans R Soc B Biol Sci* 369(1644):20130620. <https://doi.org/10.1098/rstb.2013.0620>
16. Meltzoff AN, Marshall PJ (2018) Human infant imitation as a social survival circuit. *Curr Opin Behav Sci* 24:130–136. <https://doi.org/10.1016/j.cobeha.2018.09.006>
17. Ho J, Ermon S (2016) Generative adversarial imitation learning. In *Advances in neural information processing systems*, pp 4565–4573. <https://proceedings.neurips.cc/paper/2016/file/cc7e2b878868c8bae992d1fb743995d8f-Paper.pdf>. Accessed 25 May 2022
18. Liu Y, Gupta A, Abbeel P, Levine S (2018) Imitation from observation: learning to imitate behaviors from raw video via context translation. *ArXiv170703374*. <https://doi.org/10.48550/arXiv.1707.03374>
19. Yang C, Ma X, Huang W, Sun F, Liu H, Huang J, et al (2019) Imitation learning from observations by minimizing inverse dynamics disagreement. *ArXiv191004417*. <https://doi.org/10.48550/arXiv.1910.04417>
20. Boucenna S, Anzalone S, Tilmont E, Cohen D, Chetouani M (2014) Learning of social signatures through imitation game between a robot and a human partner. *IEEE Trans Auton Ment Dev* 6(3):213–225. <https://doi.org/10.1109/TAMD.2014.2319861>
21. Boucenna S, Gaussier P, Andry P, Hafemeister L (2014) A robot learns the facial expressions recognition and face/non-face discrimination through an imitation game. *Int J Soc Robot* 6(4):633–652. <https://doi.org/10.1007/s12369-014-0245-z>
22. Boucenna S, Cohen D, Meltzoff AN, Gaussier P, Chetouani M (2016) Robots learn to recognize individuals from imitative encounters with people and avatars. *Sci Rep* 6(1):1–10. <https://doi.org/10.1038/srep19908>
23. Billard A, Grollman D (2012) Imitation learning in robots. In: Seel NM (ed) *Encyclopedia of the sciences of learning*. Springer, Boston, pp 1494–6. https://doi.org/10.1007/978-1-4419-1428-6_758
24. Billard A, Grollman D (2013) Robot learning by demonstration. *Scholarpedia* 8(12):3824. <https://doi.org/10.4249/scholarpedia.3824>
25. Breazeal C, Scassellati B (2002) Challenges in building robots that imitate people. In: Dautenhahn K, Nehaniv CL (eds) *Imitation in animals and artifacts*. Boston Review, Boston, pp 363–90
26. Nehaniv CL, Dautenhahn K (2002) The correspondence problem. In: Dautenhahn K, Nehaniv CL (eds) *Imitation in animals and artifacts*. Boston Review, Boston, pp 41–61. <https://doi.org/10.7551/mitpress/3676.003.0003>
27. Mills CM (2013) Knowing when to doubt: developing a critical stance when learning from others. *Dev Psychol* 49(3):404. <https://doi.org/10.1037/a0029500>
28. Harris PL, Koenig MA, Corriveau KH, Jaswal VK (2018) Cognitive foundations of learning from testimony. *Annu Rev Psychol* 69:251–273. <https://doi.org/10.1146/annurev-psych-122216-011710>
29. Poulin-Dubois D, Brosseau-Liard P (2016) The developmental origins of selective social learning. *Curr Dir Psychol Sci* 25(1):60–64. <https://doi.org/10.1177/0963721415613962>
30. Koenig MA, Sabbagh MA (2013) Selective social learning: new perspectives on learning from others. *Dev Psychol* 49(3):399. <https://doi.org/10.1037/a0031619>
31. Ste-Marie DM, Law B, Rymal AM, Jenny O, Hall C, McCullagh P (2012) Observation interventions for motor skill learning and performance: an applied model for the use of observation. *Int Rev Sport Exerc Psychol* 5(2):145–176. <https://doi.org/10.1080/1750984X.2012.665076>
32. Ste-Marie DM, Lelievre N, St Germain L (2020) Revisiting the applied model for the use of observation: a review of articles spanning 2011–2018. *Res Q Exerc Sport* 91(4):594–617. <https://doi.org/10.1080/02701367.2019.1693489>
33. Tomasello M, Kruger AC, Ratner HH (1993) Cultural learning. *Behav Brain Sci* 16(3):495–511. <https://doi.org/10.1017/S0140525X0003123X>
34. Tomasello M (2016) Cultural learning redux. *Child Dev* 87(3):643–653. <https://doi.org/10.1111/cdev.12499>
35. Henrich J, Broesch J (2011) On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Philos Trans R Soc Lond B Biol Sci* 366(1567):1139–1148. <https://doi.org/10.1098/rstb.2010.0323>
36. Rohbanfard H, Proteau L (2011) Learning through observation: a combination of expert and novice models favors learning. *Exp Brain Res* 215(3–4):183–197. <https://doi.org/10.1007/s00221-011-2882-x>
37. Schunk DH, Usher EL (2012) *Social cognitive theory and motivation. The Oxford handbook of human motivation*. Oxford University Press, New York, pp 13–27. <https://doi.org/10.1093/oxfordhb/9780190666453.013.2>
38. Williamson RA, Meltzoff AN, Markman EM (2008) Prior experiences and perceived efficacy influence 3-year-olds' imitation. *Dev Psychol* 44(1):275–285. <https://doi.org/10.1037/0012-1649.44.1.275>
39. Schunk DH (1987) Peer models and children's behavioral change. *Rev Educ Res* 57(2):149–174. <https://doi.org/10.3102/00346543057002149>
40. Schunk DH (1991) Self-efficacy and academic motivation. *Educ Psychol* 26(3–4):207–231. <https://doi.org/10.1080/00461520.1991.9653133>
41. Zimmerman BJ, Kitsantas A (2002) Acquiring writing revision and self-regulatory skill through observation and emulation. *J Educ Psychol* 94(4):660–668. <https://doi.org/10.1037/0022-0663.94.4.660>
42. Kirtay M, Vannucci L, Falotico E, Oztop E, Laschi C (2016). Sequential decision making based on emergent emotion for a humanoid robot. In 2016 IEEE-RAS 16th international conference on humanoid robots (Humanoids). IEEE. pp 1101–1106. Doi: <https://doi.org/10.1109/HUMANOIDS.2016.7803408>
43. Kirtay M, Vannucci L, Albanese U, Laschi C, Oztop E, Falotico E (2019) Emotion as an emergent phenomenon of the neurocomputational energy regulation mechanism of a cognitive agent in a decision-making task. *Adapt Behav* 29(1):55–71. <https://doi.org/10.1177/1059712319880649>
44. Nguyen SM, Oudeyer P-Y (2012) Active choice of teachers, learning strategies and goals for a socially guided intrinsic motivation learner. *Paladyn J Behav Robot* 3(3):136–146. <https://doi.org/10.2478/s13230-013-0110-z>
45. Hancock PA, Billings DR, Schaefer KE, Chen JYC, de Visser EJ, Parasuraman R (2011) A meta-analysis of factors affecting trust in human-robot interaction. *Hum Factors J Hum Factors Ergon Soc* 53(5):517–527. <https://doi.org/10.1177/0018720811417254>
46. Khavas ZR, Ahmadzadeh R, Robinette P (2020) Modeling trust in human-robot interaction: a survey. *arXiv:2011.04796 [Cs]*. <https://doi.org/10.48550/arXiv.2011.04796>
47. Kirtay M, Oztop E, Asada M, Hafner VV (2021) Trust me! I am a robot: an affective computational account of scaffolding in robot-robot interaction. In 30th IEEE international conference on robot and human interactive communication (RO-MAN). pp 189–196. Doi: <https://doi.org/10.1109/RO-MAN50785.2021.9515494>

48. Zonca J, Folsø A, Sciutti A (2021) If you trust me, I will trust you: the role of reciprocity in human-robot trust. *arXiv:2106.14832* [cs.RO]. <https://doi.org/10.48550/arXiv.2106.14832>
49. Kirtay M, Oztop E, Asada M, Hafner VV (2021). Modeling robot trust based on emergent emotion in an interactive task. In 2021 IEEE international conference on development and learning (ICDL). IEEE. pp 1–8. Doi: <https://doi.org/10.1109/ICDL49984.2021.9515645>
50. Kirtay M, Oztop E, Kuhlen AK, Asada M, Hafner VV (2022a) Forming robot trust in heterogeneous agents during a multimodal interactive game. In: 2022 IEEE 12th International Conference on Development and Learning (ICDL)
51. Kirtay M, Oztop E, Kuhlen AK, Asada M, Hafner VV (2022b) Trustworthiness assessment in multimodal human-robot interaction based on cognitive load. In: 2022 IEEE 31th IEEE International Conference on Robot and Human Interactive Communication (ROMAN)
52. Chen M, Nikolaidis S, Soh H, Hsu D Srinivasa S. (2018) Planning with trust for human-robot collaboration. In Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction, pp 307–315. Doi: <https://doi.org/10.1145/3171221.3171264>
53. Patacchiola M, Cangelosi A (2022) A developmental cognitive architecture for trust and theory of mind in humanoid robots. *IEEE Trans Cybern* 52(3):1947–1959. <https://doi.org/10.1109/TCYB.2020.3002892>
54. Lee Y, Hu ES, Yang Z, Lim JJ (2019) To Follow or not to follow: selective imitation learning from observations. *ArXiv191207670*. <https://doi.org/10.48550/arXiv.1912.07670>
55. Oudeyer P-Y, Kaplan F, Hafner VV (2007) Intrinsic motivation systems for autonomous mental development. *IEEE Trans Evol Comput* 11(2):265–286. <https://doi.org/10.1109/TEVC.2006.890271>
56. Bandura A (1977) *Social learning theory*. Prentice-Hall, Englewood Cliffs. <https://doi.org/10.1177/105960117700200317>
57. Bandura A (1986) *Social foundations of thought and action: a social cognitive theory*. Prentice-Hall, Inc., Englewood Cliffs. <https://doi.org/10.4135/9781446221129.n6>
58. Flake JK, Barron KE, Hulleman C, McCoach BD, Welsh ME (2015) Measuring cost: the forgotten component of expectancy-value theory. *Contemp Educ Psychol* 41:232–244. <https://doi.org/10.1016/j.cedpsych.2015.03.002>
59. Wigfield A, Eccles JS (2020) Chapter Five-35 years of research on students' subjective task values and motivation: a look back and a look forward. In: Elliot AJ (ed) *Advances in motivation science*, pp 161–98. Doi: <https://doi.org/10.1016/bs.adms.2019.05.002>
60. Frömer R, Lin H, Dean Wolf CK, Inzlicht M, Shenhav A (2021) Expectations of reward and efficacy guide cognitive control allocation. *Nat Commun* 12(1):1030. <https://doi.org/10.1038/s41467-021-21315-z>
61. Brewer SS (2008) Rencontre avec Albert Bandura: L'homme et le scientifique. [Meet Albert Bandura: the man and the scholar.]. *Orientat Sc Prof* 37(1):29–56
62. Studer B, Knecht S (2016) Motivation: what have we learned and what is still missing? *Prog Brain Res* 229:441–450. <https://doi.org/10.1016/bs.pbr.2016.07.001>
63. Eccles JS (2005) Subjective task value and the Eccles et al. model of achievement-related choices. In *Handbook of competence and motivation*. New York: Guilford Publications, pp 105–21
64. Székely M, Michael J (2021) The sense of effort: a cost-benefit theory of the phenomenology of mental effort. *Rev Philos Psychol* 12(4):889–904. <https://doi.org/10.1007/s13164-020-00512-7>
65. Friedman NP, Miyake A (2004) The relations among inhibition and interference control functions: a latent-variable analysis. *J Exp Psychol Gen* 133(1):101–135. <https://doi.org/10.1037/0096-3445.133.1.101>
66. MacLeod CM (1991) Half a century of research on the Stroop effect: an integrative review. *Psychol Bull* 109(2):163–203. <https://doi.org/10.1037/0033-2909.109.2.163>
67. Tiego J, Testa R, Bellgrove MA, Pantelis C, Whittle S (2018) A hierarchical model of inhibitory control. *Front Psychol* 9:1339. <https://doi.org/10.3389/fpsyg.2018.01339>
68. van Moorselaar D, Slagter HA (2020) Inhibition in selective attention. *Ann N Y Acad Sci* 1464(1):204–221. <https://doi.org/10.1111/nyas.14304>
69. Sidarus N, Palminteri S, Chambon V (2019) Cost-benefit trade-offs in decision-making and learning. *PLOS Comput Biol* 15(9):e1007326. <https://doi.org/10.1371/journal.pcbi.1007326>
70. Triesch J (2013) Imitation learning based on an intrinsic motivation mechanism for efficient coding. *Front Psychol* 4:800. <https://doi.org/10.3389/fpsyg.2013.00800>
71. Bandura A (1993) Perceived self-efficacy in cognitive development and functioning. *Educ Psychol* 28(2):117–148. https://doi.org/10.1207/s15326985ep2802_3
72. Belpaeme T, Kennedy J, Ramachandran A, Scassellati B, Tanaka F (2018) Social robots for education: a review. *Sci Robot* 3(21):eaat5954. <https://doi.org/10.1126/scirobotics.aat5954>
73. Mohammad Y, Nishida T (2012) Fluid imitation.-discovering what to imitate? *Int J Soc Robot* 4(4):369–82. <https://doi.org/10.1007/s12369-012-0153-z>
74. Bandura A, Ross D, Ross SA (1963) Vicarious reinforcement and imitative learning. *Psychol Sci Public Interest* 67(6):601–607. <https://doi.org/10.1037/h0045550>
75. Lowe R, Almér A, Gander P, Balkenius C (2019) Vicarious value learning and inference in human-human and human-robot interaction. In 8th international conference on affective computing and intelligent interaction workshops and demos (ACIIW), Cambridge, UK, pp 395–400. <https://doi.org/10.1109/ACIIW.2019.8925235>
76. Grollman DH, Billard AG (2012) Robot learning from failed demonstrations. *Int J Soc Robot* 4(4):331–342. <https://doi.org/10.1007/s12369-012-0161-z>
77. Saraiva M, Ayanoğlu H, Özcan B (2019) Emotional design and human-robot interaction. In Ayanoğlu H, Duarte E (eds) *Emotional design in human-robot interaction*. Human-Computer Interaction Series. Springer, Cham. pp 119–141. Doi: https://doi.org/10.1007/978-3-319-96722-6_8
78. Noroozi F, Kaminska D, Corneanu C, Sapinski T, Escalera S, Anbarjafari G (2018) Survey on emotional body gesture recognition. In *IEEE transactions on affective computing*. <https://doi.org/10.48550/arXiv.1801.07481>
79. Asada M (2015) Towards artificial empathy. *Int J Soc Robotics* 7:19–33. <https://doi.org/10.1007/s12369-014-0253-z>
80. Paiva A, Leite I, Boukricha H, Wachsmuth I (2017) Empathy in virtual agents and robots: a survey. *ACM Trans Interact Intell Syst (TiIS)* 7(3):1–40. <https://doi.org/10.1145/2912150>
81. Yalçın ÖN, DiPaola S (2019) Modeling empathy: building a link between affective and cognitive processes. *Artif Intell Rev* 53:2983–3006. <https://doi.org/10.1007/s10462-019-09753-0>

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

Johann Chevalère is a postdoctoral researcher at the University of Potsdam and the Cluster of Excellence “Science of Intelligence”, Berlin. He received a PhD degree in cognitive psychology from the University of Bordeaux (France) in 2014 and was a postdoctoral researcher in the Department of Psychology at the University of Fribourg (Switzerland) in 2017 and at the Laboratoire de Psychologie Sociale et Cognitive, CNRS and Clermont-Auvergne University (France) in 2018. His research interests focus on the interactions between digital technologies for education and learners’ social, emotional and cognitive characteristics.

Murat Kirtay is an Assistant Professor at the Department of Cognitive Science and Artificial Intelligence, Tilburg University. He received a PhD. degree in BioRobotics from the BioRobotics Institute of the Scuola Superiore Sant’Anna (Pisa, Italy) in 2019. He was a postdoctoral researcher at Adaptive Systems Group at Humboldt-Universität zu Berlin in Germany, where he conducted interdisciplinary research at the intersection of robotics, neuroscience, and psychology. His research interests span humanoid robotics, computational neuroscience, cognitive systems, and machine learning.

Verena Hafner is Professor of Adaptive Systems at Humboldt-Universität zu Berlin and Head of the Adaptive Systems Group at the Department of Computer Science. She holds a Master with distinction in Computer Science and AI from the University of Sussex, UK, and a PhD from the Artificial Intelligence Lab, University of Zurich, Switzerland. Before moving to Berlin, she worked as an associate researcher in the Developmental Robotics Group at Sony Computer Science Labs in Paris, France. She is part of the Programme Committee of the DFG Priority Programme The Active Self (SPP 2134), PI in the DFG Cluster of Excellence “Science of Intelligence”, PI in the EU H2020 project ROMI on Robotics for Microfarms (2017-2022) and PI in the EU H2020 project HIVEOPOLIS (2019-2024). She is reviewing for the EU, BMBF, DFG, AvH, journals and conferences. Her research interests include sensorimotor interaction and learning, joint attention, internal models, and exploration strategies.

Rebecca Lazarides is an educational researcher and full Professor for Research on School and Instructional Development at University of Potsdam, main-PI in the Cluster of Excellence “Science of Intelligence”, member of the advisory board of the Ministry for Education, Youth and Sports in the federal state Brandenburg / Germany, member of the advisory board of the Institute for School Development of the Länder Berlin and Brandenburg (ISQ), and Editorial Board Member of the scientific peer-reviewed journals Learning and Instruction and the German Journal for Educational Research (ZfE). Her research interests include learning and instruction, motivation and emotion, interindividual motivational differences, teacher professional competence and adaptive AI in education.