Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

"Teach Me About Objects!" – Experience-Driven Interaction for Teachable Robots

Tuan Vu Pham, Chao Wang, Heiko Wersing, Marc Hassenzahl

2025

Preprint:

Copyright ACM 2025. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in ACM Designing Interactive Systems Conference (DIS) 2025, https://doi.org/10.1145/3715336.3735811.



"Teach Me About Objects!" – Experience-Driven Interaction for Teachable Robots

Tuan Vu Pham
Ubiquitous Design / Experience & Interaction
University of Siegen
Siegen, North-Rhine Westphalia, Germany
Honda Research Institute EU GmbH
Offenbach am Main, Germany
tuan2.pham@uni-siegen.de

Chao Wang
Honda Research Institute Europe GmbH
Offenbach am Main, Germany
chao.wang@honda-ri.de

Heiko Wersing
Honda Research Institute Europe GmbH
Offenbach am Main, Germany
heiko.wersing@honda-ri.de

Marc Hassenzahl
Ubiquitous Design / Experience & Interaction
University of Siegen
Siegen, North-Rhine Westphalia, Germany
marc.hassenzahl@uni-siegen.de

Abstract

To adapt to specific places and people, robots must recognize objects, which are typically taught by users—a tedious process. Inspired by anecdotes of positive teaching experiences shared by educators, sports coaches, and animal trainers, we developed seven experience-driven ways to make teaching a robot more engaging. For example, one interaction involved the robot prompting users to tell personal stories about the objects. A video vignette study (N=184) showed that experience-driven teaching was perceived as more positive than the current technology-driven teaching. Participants reported feeling more competent, connected to the robot, and valued. Additionally, the robot was perceived as more extroverted, open, agreeable, and conscientious. Overall, the experience-driven design of teaching interactions enhanced engagement and persistence by fostering reciprocal exchange and mutual understanding. In addition, the study lends support to an anecdotal approach to designing positive experiences through technology.

CCS Concepts

• **Human-centered computing** → **Empirical studies in interaction design**.

Keywords

Human-Robot Interaction, Teachable Robots, Adaptive Robotics, Interactive Teaching Strategies, Object Teaching, Machine Learning, Supervised Learning, Computer Vision

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

DIS '25, Funchal, Portugal

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-1485-6/25/07
https://doi.org/10.1145/3715336.3735811

ACM Reference Format:
Tuan Vu Pham, Chao Wang, Heiko Wersing, and Marc Hassenzahl. 2025. "Teach Me About Objects!" – Experience-Driven Interaction for Teachable Robots. In *Designing Interactive Systems Conference (DIS '25), July 05–09, 2025, Funchal, Portugal.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3715336.3735811

1 Introduction

Robots are physical machines designed to interact autonomously with things, people, and places. To do this, they need to learn. While some of this learning can be self-directed and implicit, many situations require human assistance. For example, for a robot to fetch a favorite mug from the kitchen, the owner has to teach the robot which mug this is. One way for laypeople to teach robots is *programming by demonstration* or *imitation learning* [4, 8, 14]. In the case of object learning, robots require a human to show the object (e.g., a mug) and to link it to relevant concepts (e.g., mug, favorite, container).

So far, work on teachable robots has focused primarily on the capabilities and informational needs of the robot. While this is certainly important, people must also be and remain motivated to teach robots. Unfortunately, teaching a robot is currently a tedious task that humans are reluctant to perform [40]. This is due to an incongruence between how people prefer to teach and how robots are programmed to learn [36]. It seems likely that without establishing enjoyable and engaging teaching practices, people will not engage in teaching robots, which in turn will impede the robot's capabilities due to fewer contextual data. The consequence may be a loop of progressively worsening motivation to teach, resulting in a lack of data acquisition and insufficient robot performance [7, 34]. Furthermore, tedious teaching leads to a negative user experience, frustration, and the eventual abandonment of an otherwise capable robot [2]—an unsustainable struggle.

Future robot users in certain settings, such as domestic environments, informal and formal care, may be particularly affected, as they may not be familiar with how robots acquire and process information [11]. To counteract this, we propose to enrich teaching

robots with a focus on the potential psychological and emotional needs of the human teacher.

The present paper explores the design of engaging ways for people to teach in-home robots to recognize particular objects. We start from the existing, technology-driven teaching interaction of a widespread mobile robotic platform that possesses fully realized object-learning and recognition capabilities (*Johnny*, a customized *Scitos G5* robot by *Metralabs*, see Section 3.2).

Using a wellbeing- and experience-driven design approach [22, 25], we first conducted semi-structured interviews with educators and trainers to gather anecdotes of positive (i.e., enjoyable, meaningful) teaching experiences and practices. Based on these positive practices, we developed seven experience-driven object-teaching interactions and realized them as video vignettes. In an online experimental vignette study (N=184) [1, 6], the experience-driven teaching interactions were tested against the technology-driven interaction currently implemented in the robot *Johnny*. The focus was on the quality of the interaction perceived by the participants and the resulting willingness to engage in the teaching. The insights contribute to a more experience-driven way of designing teaching interactions for continually learning robots [7, 12]. We will first discuss prior work on teaching robots. We will then elaborate on the theoretical foundations of our design approach and present the experience-driven interaction designs based on the anecdotes gathered through the interviews with educators. Subsequently, we present a video vignette study to assess the designs. Finally, we consider the implications of these findings for human-robot interaction research and design.

2 Teaching Robots – Approaches for Non-Expert Users

With computer vision and machine learning, robots can learn to recognize and name objects based on their camera-based sensors. Although they can also be trained from web-based image collections [12], specific learning examples from the user allow the robots to adapt to complex and chaotic environments, such as a user's home. This is in contrast to learning for more standardized and organized environments, for example, for industrial robots on factory floors [12, 35]. The context in which both the robot and its teacher are situated matters, as it directly impacts the appropriate learning and teaching strategies. As a result, what is needed are sufficiently intuitive ways for home users with limited technical expertise to teach a robot [4, 29, 43].

A central finding is that humans tend to perform the teaching differently than technically required by most robots [34, 40, 41]. Humans rely on established teaching practices that include feedback, the constant assessment of the learner's knowledge and understanding, as well as the provision of socio-emotional care. In other words, human teachers want to give feedback to the robot, motivate it, and adapt their teaching behavior based on their understanding of the robot's learning procedures. This suggests that the robot should transparently communicate its level of understanding to the teacher, e.g. "by demonstrating their current knowledge and mastery" of the skill of concern [38]. Humans think of learning as reciprocal and social. This is in stark contrast with traditional supervised learning for robots, where the machine receives an often heavily standardized input, processes the information in a black box that is incomprehensible to the human teacher, and delivers an output without further explanation. While this is especially true for the broad variety of industrial robots, available domestic robot learners have yet to adopt learning principles that differ significantly from traditional models and provide more socially intuitive teaching interactions for non-expert users.

By building upon insights from psychology, pedagogy, developmental sciences, and sociology, Calinon and Billard [11] suggested that it would be important to better understand the human teacher's role in the robot's learning loop to design efficient systems. A literature review by Vollmer and Schillingmann [40] makes a similar claim and reveals the overall scarce number of studies conducted on the topic [35, 38]: Only 18 papers so far considered the human teacher's role and behavior. They concluded that humans are more open towards adapting their teaching style to an active learner, which approaches them for knowledge acquisition, rather than a passive one. The same applies to feedback: The robot's response greatly impacts the teacher's idea of how the teaching content is taken up and how it needs to be adjusted. To that end, the robot's gaze, gestures, or additional mediating artifacts have been used to better convey the robot's perception and progress. For example, Rouanet and colleagues [34] found that mediating tools (e.g., a smartphone that displays the robot's field of vision or a pointing device that controls the robot's line of sight) helped to guide the robot's point of focus in an effortless and entertaining way for humans when performing an object-teaching task. The focus guidance can also be dealt with by allowing users to hold an object in front of the robot's camera sensor. The corresponding video feed is then shown on a display mounted on the robot, acting as a feedback mechanism. The human teacher can then immediately adapt and manipulate the object to cover challenging viewing angles. This greatly reduced the amount of additional teaching turns necessary [45], leading to quicker teaching successes.

A recent study by Ayub and colleagues [7] expanded this by examining how human teachers' behavior changes when teaching a continuously learning robot across multiple sessions. They observed that the initial novelty and excitement of interacting with the robot waned over time. To counteract such disengagement, an earlier work attempted to make teaching robots more engaging by broadly likening it to human- and especially children-learning, for example, by letting the robot ask simple or slightly intricate follow-up questions on a scanned object's features to elicit a better teacher response. This was based on the assumption that users simply want to teach a robot, due to humans' 'inherent ability' to impart knowledge to their offspring and the desire to witness the learner's progress [10]. However, teachable robots have so far provided little reason to be met with a comparable emotional response. The main takeaway is that prior work focused on physiological and cognitive ergonomics to achieve computational improvements, sidelining research on user experience [2], as reflected in Vollmer's literature review, where effective robot performance was prioritized over human teacher experience [40].

Interestingly, Rouanet and colleagues noted in passing in their study of 'facilitation tools' that embedding the "still [...] unusual and artificial task" of repeatedly presenting and teaching objects to a robot in a somewhat plausible and entertaining game context

could help maintain the user's interest and teaching motivation [34].

Elaborating on this, we set out to design experience-driven teaching interactions that foster emotional involvement with a domestic robot learner by building on positive moments from real-life teaching experiences to make teaching interactions emotionally resonant and relatable. These seven interactions have in common that they portray the robot as eager to learn, willing to share its perceptions, or wanting to associate objects with the teacher's personal routines. We do this in a way that is not deceiving but acknowledges the robot's machine-ness and unique learner role.

We suggest that the tediousness and artificiality of teaching a robot is not only due to the predictable principles that govern its system but also to the limited consideration of socio-psychological factors in human-robot interaction (HRI).

3 Positive Teaching Interactions for Learning Robots

3.1 Collecting Positive Teaching Practices

To make the experience of teaching of objects to robots more enjoyable and engaging for humans requires an understanding of what constitutes a "positive" (i.e., gratifying and meaningful) teaching experience. In line with Klapperich and colleagues' approach ([25], see also [22]), we gathered anecdotes of positive teaching experiences in various contexts, which serve as an inspiration for the later design of the experience-driven human-robot teaching practices.

We relied on interviews with educators and coaches rather than observations (e.g., attending a class or soccer practice) since only interviews provide access to internal thought processes, motivations, and meanings surrounding positive moments in teaching. While meaningful teaching also occurs in informal, non-professional contexts — such as a parent guiding a child or a peer helping a colleague— we focused on professional teachers and coaches for this exploratory phase because of their ability to reflect deeply on years of interactions with a wide variety of learners and to articulate clearly what makes teaching experiences personally enjoyable and motivating. This approach yielded a rich collection of anecdotes, encompassing a broad spectrum of experiences, which provided valuable insights into the aspects and mechanisms that contribute to distinctly positive teaching moments.

While the anecdotal approach can be further backed by existing empirical and theoretical insights from pedagogy and psychology, it also addresses a notable gap in current research. Previous reviews—e.g., Sinclair [37] and more recently by Han and Yin [18]—identify several factors that influence teacher motivation, including relationships with students, student motivation and attitudes, and the satisfaction derived from imparting knowledge. However, such research remains somewhat vague regarding how these factors manifest in specific, moment-to-moment dynamics.

To address this gap, we gathered a series of firsthand teaching anecdotes, focusing on specific, positive teacher-student interactions that help create meaningful teacher motivation. This approach complements existing research by grounding abstract motivational factors in concrete, lived experiences.

All in all, we conducted six semi-structured interviews with schoolteachers, university lecturers, sports coaches, and a dog trainer. Interviewees were recruited through referrals (2 male, 4 female, age range 28 to 62, see Table 1 for an overview). The interviews were held via Skype and recorded with the consent of the interviewees. The average duration of the video calls was 55 minutes.

Table 1: Demographic data for the interview partners.

Educator/Instructor Type	Count	Gender	Age
Dog Trainer	1	f	62
School Teachers	2	f, f	28, 35
Sports Coaches	2	m, m	31, 35
University Lecturer	1	f	50

Our questions focused on positive moments encountered while teaching others. For example, we asked "What made you take on teaching?", "What situations gave you a sense of accomplishment?", "How does a learner best signal to you that they need help/are understanding?", "What motivates you to support an individual learner to achieve certain goals?". Often, the responses consisted of accounts of specific situations experienced in a particular teaching context. To elicit additional details, we asked follow-up questions on the fly that related to the anecdotes being told.

The interview transcripts were analyzed by the first author using thematic analysis [9]. After a systematic review of the transcripts to generate codes and themes, a second coder independently reviewed and labeled the transcripts based on the initial code list. The resulting themes were then compared and refined by both coders, with an emphasis on rich descriptions of standing-out moments. Positive teaching practices were formulated by including further abstract descriptions of the circumstances, elements, activities, and aspects central to the specific instances of teacher-learner exchanges (see Table 2 for an overview of the identified practices).

For example, the university lecturer explained that she likes to be asked specific questions regarding her course material. This gave her the impression that the learner had at least attempted to understand the subject matter before asking. The lecturer would then be better able to relate and more competently engage with the student's current understanding and help the student achieve the learning goal. This requires the teacher to be able to adapt to the learner's request. At the same time, the student needs to be self-motivated, possess basic research skills, and be able to ask questions comprehensively regarding the course material. Students should ask questions that convey their current understanding to help the lecturer explain things more concisely.

A middle school language teacher shared that she appreciated "catching" her students learning outside of class, such as during a break. This gives her the feeling that she has a positive impact on the students beyond the classroom. Not only does this validate her teaching skills, but she also feels popular, as the pupils choose her subject over other school subjects or recreational activities. For this positive moment to occur, the teacher would need to be able to recognize and appreciate the extra effort. The learner needs to concern themselves with the learned material during their own time.

As the examples suggest, the activity of teaching and the situations experienced through it are heavily intertwined with the learner's learning activities and attitude. We recognize that the reciprocity of the teaching-learning dyad is shaped not only by the teaching material and 'know-how' of both sides but also by what moves and motivates the teacher and the learner.

3.2 Current Form of Object Teaching: The Robot Johnny

The technology-driven interaction for teaching objects, which we build upon in the present paper, is currently implemented on an experimental robot called *Johnny*. It is a mobile robot based on the *Scitos G5* by *METRALABS*, developed for research on autonomous operation in less controlled, more natural environments [19]. *Johnny* features advanced object-learning and recognition capabilities, that allow it to continuously adapt by retraining its machine-learning model with visual learning examples provided by a human teacher.

The system consists of a wheeled, cylindrical base that houses two computers featuring robot navigation and deep-learning-based visual object learning. The setup has a *Kinect 2* camera and a microphone array mounted on top of a metal strut. Below this headpiece, a touch screen is mounted on the robot, facing forward (Figure 1). This configuration allows for an online learning system, in which the user presents an object to the robot's camera.

Online means that the object recognition algorithm is adjusted by the visually acquired data rather than trained from a static data set (i.e., offline learning). The system then segments, classifies, and collects relevant object features from the captured images using a convolutional neural network. Through a dialog manager that employs verbal and visual communication through speakers, microphones, and the touch display, the system communicates a predicted class for the presented object to the user. It is then up to the user to validate or correct the robot's prediction. Subsequently, a label can be provided through speech-to-text to update the corresponding object data for future reference [19]. The emerging, technology-driven teaching interaction serves as a baseline in the present paper.

3.3 From Positive Practices to Teaching Interactions

Our goal is to make the object teaching experience more engaging by enriching the technically necessary interaction with elements from the collected positive practices, thereby reframing the object learning of the robot in a plausible and more enjoyable way. The approach we use [26], suggests a six-step process: (1) to scope the design space (here: robot teaching), (2) to gather examples of especially positive practices relevant to the design space (here: positive human-human/human-animal teaching practices), (3) to consolidate positive practices to make them actionable for design (see Table 2), (4) to design ideal practices, (5) to specify the material to shape the ideal practices (here: teaching interactions), and (6) to realize the material (here: video vignettes). The general idea of this design approach is to shape meaningful practices through the particular interaction with the given technology.

The notion of a positive practice merits some further explanation. Klapperich and colleagues' [26] design approach builds upon a praxeological view of human activity [36]. In this, activities, that is, practices, consist of three key elements: know-how, meaning, and material. Know-how encompasses all the knowledge, especially the practical, necessary to perform a practice. Meaning encompasses the subjective experience of performing the practice, with a focus on underlying motivations, values, and beliefs. Finally, material encompasses all the materials (i.e., technologies) crucial to the practice. The three elements are intertwined, that is, changes in the one will likely lead to changes in the other. For example, a "smart" technology (i.e., material) may lead to changes in the know-how required to perform a practice, which may in turn impact feelings of competence gained through the practice (i.e., meaning). Through this interdependence, practices can become "units of design" [27].



Figure 1: Left: The mobile robot unit *Johnny* [42] Right: Its low-fidelity equivalent used in the video vignettes of the present study.

Following this, we broke down the positive practices from the anecdotes into three key elements that form the foundation of the teaching/learning dyad (see Table 2 for an overview of the teaching practices and interaction designs). Building on these elements, we developed seven robot teaching interactions based on the technology-driven baseline. Rather than presenting an exhaustive framework, these interactions are the result of a design exploration inspired by successful human-human and human-animal teaching practices.

3.3.1 Technology-Driven Teaching Interaction – Baseline (1). The user initiates a teaching session by calling the robot ("Hey Johnny") and starting the object learning program with a voice command ("Object learning"). The teacher presents the object to *Johnny*'s camera, which displays the video feed on its front display and highlights the detected object with a frame. To begin recording data, the user says "Record". Images are captured as the teacher turns the object to show it from various angles. This continues for

Table 2: Teaching Interactions and Interviewees' Positive Practices that inspired them.

Interaction Design	Positive Practice	Meaning (According to Interviewees)	Learner's Know-How	Teacher's Needs Addressed
Technology-Driven Baseline (1)	-	-	-	-
ProActive (2)	Learner actively approaches teacher for a lesson.	"I would like to see a spark of initiative from the students." (University Lecturer)	Indication of self-motivation	Relatedness, Competence
ColorcodeML (3), PatternML (4)	Learner shares their perception of the learned (through their own words and notions) and demands to know about the teacher's view as well.	"When training my dogs, I had to learn their language. I needed to listen to and understand what they wanted, too." (Dog Trainer)	Conveying own and asking about teacher's perception	Relatedness, Competence
Storytime (5)	Learner transforms the learned into a creative depiction of their own - for a shared discourse with the teacher.	"Pupils connect best with the taught by creatively expressing how it relates to them, through drawing and handicrafts." (Schoolteacher)	'Artistically' expressing themselves	Relatedness
RecogHabits (6)	Learner attentively relates the learned to the teacher's routine through contextual behavior.	"I love it when a dog comes to my bedside to wake me right before my alarm clock rings. I feel understood and valued." (Dog Trainer)	Contextualizing the learned with teacher's habits	Security, Relatedness, Competence, Popularity
ButlerDevelop (7)	Learner makes their relative progress apparent to the teacher.	"I like seeing lesser players on the team visibly catching up and reaching important milestones." (Sports Coach)	Presenting own progress	Relatedness, Competence, Popularity
AsleepStudy (8)	Learner indicates their will and effort to rehearse the learned beyond the initial teaching situation.	"Spotting a pupil practicing [foreign language] vocabulary by themselves during recess shows me that I encouraged them to learn." (Schoolteacher)	Visibly engaging in learning during 'own' time	Relatedness, Competence, Popularity

30 seconds until enough frames are collected or the user says, "Stop recording".

After recording, the robot processes the images for shapes and patterns, attempting to identify whether the object is already known or new. The user can confirm this or provide a label for the new object, like "Coffee Pot". The session concludes with the robot repeating the label and displaying an image of the learned object, along with similar known objects and their statistical likelihoods. Overall, this interaction narrowly focuses on the technical aspects of the object-scanning and is performance-oriented. There is little indication of a meaningful goal for the human teacher or the robot learner beyond expanding the robot's object recognition capabilities (cf. Figure 2a, Baseline (1)), enacted with robot *Johnny*'s low fidelity cardboard stand-in in the video vignette *Baseline (1)* to ensure comparability between the videos). This technology-driven interaction is fully implemented and is currently used by the robot *Johnny*. It acts as the baseline interaction in our study—as a point of reference and basic functionality template for our seven experience-driven teaching interactions.

3.3.2 Pro-Active – ProActive (2). While this concept closely resembles the baseline, it presents *Johnny* as a proactive learner: Instead of the user initiating the learning session by calling the robot, the robot actively approaches the user and asks to be taught. In addition, the robot inquires to be taught further objects after scanning the initial one (Figure 2a, ProActive(2)).

The positive practice that inspired this approach would be the university lecturer's preference for self-motivated students who attempt to gain a deeper understanding of the subject matter by eagerly posing relevant follow-up questions. Ideally, this condition

results in a greater willingness of the teacher to engage. Allowing both sides to approach one another and emphasizing the robot's role instead of having the situation dominated by the user's goals should address the need for relatedness and shared accomplishment.

3.3.3 Colorcode Machine Learning – ColorcodeML(3). After processing and labeling the image data, *Johnny* asks the user for the object's color (e.g., "the apple is red"), acknowledges the response, and saves it for future reference. The interaction continues as *Johnny* conveys its perception by providing a hexadecimal color code (e.g., "D22B07") for the object's hue and lists similarly colored objects (Figure 2b, ColorcodeML(3)). Although the color code may not be meaningful to the user, it suggests that *Johnny* has its own way of perceiving and categorizing learned objects.

This concept is inspired by positive moments from a university lecturer and a dog trainer. The lecturer explained that knowledge may mean different things to the learners, encouraging them to build their own mental models. With the learner willfully conveying their perception, the teaching process would become more personal and relatable to the lecturer. This requires teachers to be accepting of deviations from their perception. The dog trainer suggested that a successful trainer-trainee relationship relies on mutual understanding. Dog owners should not only give commands but also strive to understand the dog's needs and language, setting aside their sense of superiority to foster trust. Similarly, we believe that the robot should not hide its machine-ness or pretend to be human. By providing its perspective and clear motivations for its actions, the robot can help users better make sense of teaching and learning of objects.

3.3.4 Pattern Machine Learning – PatternML (4). The "Pattern Machine Learning" concept is similar to "Colorcode Machine Learning", but it more closely illustrates how the object classification method of the system works—based on patterns, shapes, and shades. Instead of a hexadecimal code for the object's color, *Johnny* displays objects possessing similar visual features (Figure 2c, PatternML(4)).

While this also leans toward emphasizing the otherness of the robot, it also slightly leans in the direction of explaining its actual mode of operation. Although this can be considered as a feature concerning instrumentality and usability (i.e., lessening the opaqueness of the black box), we felt this approach would help users to better relate to the learner's motivation.

3.3.5 Story Sharing – Storytime (5). After teaching an object, *Johnny* will ask for a personal story related to it. The user can share a brief anecdote, which the robot will visualize through image generation. This image is saved with the object data and referenced later when *Johnny* retrieves the object. For instance, if the object was a gift from a friend, the display would show an image of the friend, a wrapped gift, and the teacher's face (Figure 2d, Storytime (5)). Later, *Johnny* might say, "I am bringing you the coffee pot. The one you got from your friend as a gift".

As noted by a schoolteacher in our interview, allowing pupils to express learning content through artistic means—like drawing a story—helps them to better understand and connect with the material. This approach also allows teachers to empathize with their students' perspectives. With the 'creative' depiction by the robot (which can be considered going beyond the simple repetition of the taught object and story), and later references to the story, we aimed to foster a more personal connection between the teacher, the robot, and the object.

3.3.6 Recognizing Objects and Habits – RecogHabits (6). After the initial object teaching, *Johnny* observes the teacher interacting with the object at specific times. After a few days, *Johnny* begins to automatically fetch the object for the user during those times (Figure 2e, RecogHabits (6)). This idea is inspired by a dog trainer's positive experience with her dog coming to her bedside at her usual waking time, showing that the dog understands her routines. This aims to make teaching more meaningful by applying learned knowledge to the user's daily life. The user may feel closer to and valued by the robot, interpreting its actions as being considerate and supportive in their shared environment.

3.3.7 Raising a Butler Game – ButlerDevelop (7). The teaching includes game-like elements with a gauge that tracks the robot's development. Progress is made by teaching more objects, leading to new abilities like fetching specific items. A visual cue for leveling up involves a change in the robot's appearance, such as presenting the robot with a pacifier and later with a dress jacket (Figure 2f, ButlerDevelop(7)).

A soccer coach noted that witnessing players improve over time is one of the most rewarding aspects of his job. He values individual progress, which validates his training methods and fosters trust between him and the team. Therefore, indicating learner progress would be important. While gamification may not help the artificiality of the object teaching process, visualizing past and future developments creates motivation through anticipation for shared successes between the teacher and learner.

3.3.8 Studying While Asleep – AsleepStudy (8). After the initial learning interaction, *Johnny* checks the object data during its "free time", such as when standing in its charging dock at night. Instead of idling or showing a charging animation, the robot shows the objects being re-scanned visually, while audibly repeating their labels. The teacher's face and name are also displayed to reinforce the connection between the teacher's effort and learning (Figure 2g, AsleepStudy(8)). When the teacher sees *Johnny* repeating the taught objects, he may feel that he has meaningfully influenced the robot's learning progress. This is similar to a middle school teacher who reported "catching" her students studying foreign language vocabulary from her lessons during recess instead of playing or engaging in other recreational activities.

4 Online Video Vignette Study

4.1 Participants and Procedure

The video vignette study was carried out online with SurveyMonkey[1]. Participants were recruited via Prolific[2]. Each received a monetary compensation equivalent to $ 24 for their participation. The study took between 36 and 42 minutes to complete. The full sample consisted of 184 participants (89 male, 95 female, 0 nonbinary, age range 18 to 68, median = 36). Since the study was in German, participants were screened to be fluent in this language.

[1]https://www.surveymonkey.com/
[2]https://www.prolific.com/

Each participant was confronted with all eight teaching interaction video vignettes.

To evaluate the impact of the experience-driven teaching interactions on subjective experience, we conducted an experimental vignette study (EVM), a widely used method in the fields of psychology, sociology, and education [5]. Experimental vignettes, consisting of systematically varied short texts or videos, have recently gained traction in HCI research [31, 32]. This method allows consistent experimental manipulation of factors (in our case, how the teaching interaction takes place) and systematically tests the effect of the factor on the outcome variables of interest. Since it uses videos, experiments can be done without realizing the interaction designs on the robot. While less expensive, it also helps avoid technical issues with a prototype (e.g., reaction latency, inconsistent robot actuation, software glitches). It further enables broader participant samples and scenarios; for instance, we employed eight scenarios with over 180 participants to assess the robustness of possible effects. With the actual object learning robot *Johnny*, this would have required programming the seven new interactions, as well as inviting nearly two hundred people to come to the lab in person. Although the hypothetical nature of EVM poses a challenge, careful scenario depiction and participant checks address this (e.g., level of immersion when imagining themselves in the scenario). Research confirms consistent results across formats like text and images [33]. EVM's benefits in terms of resources, participant numbers, and control are particularly valuable for early-stage research.

After a short introduction, each participant was presented with eight short video vignettes (with a duration of one minute on average). Each video vignette depicts the object teaching interaction between a human user and the robot *Johnny* represented by a low-fidelity prototype of the mobile robot (Figure 1). The text-to-speech engine *Balabolka*[3] was used to record lines in a male synthetic voice that matched *Johnny*'s voice. The graphical user interfaces (cf. Figure 1, Baseline (1)) are based on screens of the actual robot. In the vignettes, the teaching takes place at a dinner table in a home kitchen. While the basic object teaching mechanism remained similar in all videos, the way the interaction played out varied between vignettes. After each clip, participants were asked to imagine being the teacher and rated how they would experience the (imagined) interaction with the robot. This included questions on need satisfaction, positive affect, fun, perceived learner personality traits, and technology relationship felt during the teaching activity (see section "Measures"). The videos were shown in a random order. The survey concluded after all eight vignettes were assessed. Assuming a sufficiently large sample size, randomization allowed us to accomplish several things: We were able to determine attitudes and ratings for each vignette with relatively high reliability and validity. We can also use data from the few people who dropped out of the study before viewing the full set [3, 15]. Since we also show how the teaching is currently implemented, we will get an indication of how user perceptions vary relative to this baseline.

[3]https://www.cross-plus-a.com/balabolka.htm

4.2 Measures

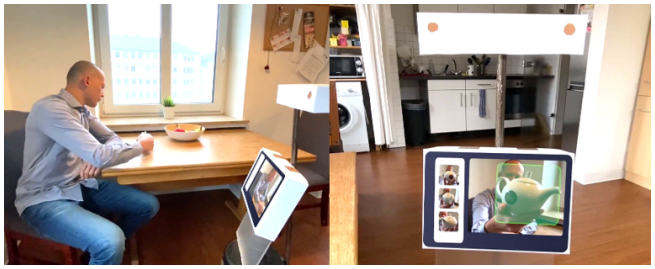
Following previous work (e.g., [21, 23]), we operationalized the quality of the teaching experience by measuring the user's experience. User experience is conceptualized as a primarily affective response to using a technology in a particular practice (here: teaching). Whether this response is positive (i.e., enjoyable and/or meaningful) is a consequence of the fulfillment of psychological needs (i.e., autonomy, competence, relatedness, stimulation, popularity, physicality, and security). From our interviews with educators, we identified four key needs as important for positive teaching practices, namely competence, relatedness, popularity, and security [13, 21] (see Table 2). The needs autonomy and physicality seemed less relevant given that the present teaching interaction focuses on social aspects. Each of the four needs was measured with two items. Examples are "I felt that I was completing the task of teaching the robot successfully" (competence) or "I felt close and connected with the robot" (relatedness) (see Table A.1 in the Appendix for the complete set of questionnaire items). We computed overall *Need Satisfaction* (NS) by averaging the responses to the eight items per vignette (i.e., within-subjects design).

Participants responded to the questionnaire items on a 7-point intensity scale ranging from 1 = "not at all" to 7 = "extremely". We used this response form for all scales employed in the present study. Table 3 shows the internal consistency (Cronbach's alpha) and discriminant validity (as intercorrelations, Pearsons's r) for each scale on the pooled data. For instance, the internal consistency for the NS scale was very high (Cronbach's α = .92).

Positive affect (PA) was measured with the Positive Affect Negative Affect Schedule [44]—in particular, we chose the short-form as a psychometrically acceptable alternative (I-PANAS-SF [39]). The PA scale consists of the five adjectives active, determined, attentive, inspired, and alert. We forewent measuring negative affect due to the extensive length of our survey and our research focus on positive emotions invoked by the interactions presented. The internal consistency of the PA scale was high (Cronbach'sα = .93). We calculated the *Positive Affect* (PA) as the average mean of the five items for further analysis.

In addition, we constructed the *Teaching Fun* scale to assess participants' enjoyment and engagement when interacting with the robot during the object teaching. The scale included the following items: "Teaching the robot gave me pleasure" and "Teaching the robot amused me". The internal consistency of the *Teaching Fun* scale was high (Cronbach's α = .91). We calculated the average mean of the two items for further analysis.

We further constructed scales based on the Big Five personality traits: extraversion, openness, agreeableness, conscientiousness, and neuroticism [16] to measure the robot's perceived personality. Except for neuroticism, higher trait scores are associated with better academic performances [17, 30]. The five personality traits were each measured with two items. Typical items were "I felt the robot was open to learning new things" (openness), "I felt the robot was very cooperative and forthcoming" (agreeableness), or "I felt the robot was paying attention to detail" (conscientiousness) (see Table A.1 in Appendix for all items). The reliability analysis indicated good internal consistency of extraversion, openness, agreeableness, and conscientiousness. Neuroticism will be excluded from further



(a) Baseline (1) and ProActive (2)



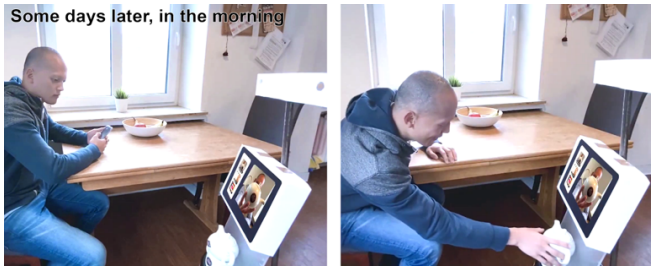
(b) ColorcodeML (3)



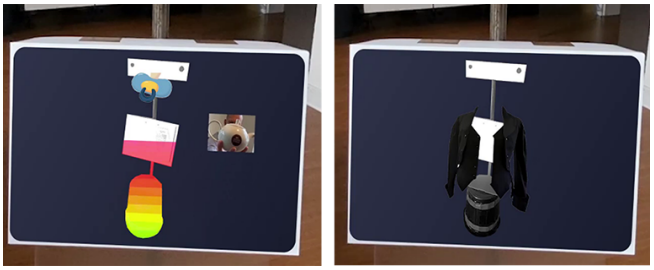
(c) PatternML (4)



(d) Storytime (5)



(e) RecogHabits (6)



(f) ButlerDevelop (7)



(g) AsleepStudy (8)

Figure 2: Overview of the one technical teaching interaction Baseline (1) previously used with the object-learning robot Johnny and the seven new experience-driven teaching interactions ProActive (2), ColorcodeML (3), PatternML (4), Storytime (5), RecogHabits (6), ButlerDevelop (7), and AsleepStudy (8).

analysis due to its rather low internal consistency (α = .64), which suggests it may not be as reliably measured in this context. For each personality trait, a scale value was computed by averaging the respective items per vignette. The intercorrelations among the personality traits extraversion, openness, agreeableness, and conscientiousness are high and, in some cases, close to Cronbach's Alpha (see Table 3). This implies strong relationships among the four traits in the sense of a more or less sociable personality overall.

Table 3: Internal consistency (Cronbach's Alpha) and discriminant validity (intercorrelations, Pearson's r) (pooled data, N=1179). NS = Need Satisfaction, PA = Positive Affect, Fun = Teaching Fun, Ext = Extraversion, Ope = Openness, Agr = Agreeableness, Con = Conscientiousness, Neu = Neuroticism, Alt = Alterity, Imm = Immersion.

Measure	NS	PA	Fun	Ext	Ope	Agr	Con	Neu	Alt	Imm
Need Satisfaction	(.92)									
Positive Affect	0.89	(.93)								
Teaching Fun	0.70	0.72	(.91)							
Extraversion	0.78	0.79	0.66	(.84)						
Openness	0.59	0.56	0.70	0.74	(.90)					
Agreeableness	0.73	0.70	0.75	0.83	0.82	(.85)				
Conscientiousness	0.71	0.68	0.76	0.85	0.84	0.92	(.83)			
Neuroticism	0.38	0.41	0.40	0.35	0.11	0.21	0.20	(.64)		
Alterity	0.67	0.65	0.70	0.77	0.69	0.80	0.79	0.36	(.71)	
Immersion	0.61	0.61	0.70	0.61	0.64	0.69	0.65	0.09	0.58	(.91)

Finally, participants assessed the perceived *Alterity relationship* with the robot. This was conceptualized as a unipolar dimension of the feeling that the robot is an interacting counterpart or autonomous actor (see [20, 24, 28]). Examples of items that measured alterity were "I felt that the robot was a discrete thing separate from me" and "I felt that I acted in cooperation with the robot" (see Table A.1 in Appendix for all items). The internal consistency for the alterity scale was acceptable (Cronbach's α = .71). A scale value for alterity was computed by averaging the four items per vignette.

To check for confounding factors, participants additionally rated for each video whether they could imagine themselves being in the situation shown ("I was able to put myself into the situation shown in the video" and "I could relate to the scenario presented in the video"). The mean of the two items was calculated and used as a scale for subsequent analysis. The internal consistency of the immersion scale was high (Cronbach's α = .91).

4.3 Hypotheses

We designed seven new experience-driven teaching interactions and compared them to the technology-driven baseline interaction currently used by the *Johnny* robot. We expect the following:

- **H1:** Experience-driven teaching interactions will result in higher psychological need satisfaction compared to the technology-driven baseline interaction.
- **H2:** Experience-driven teaching interactions will elicit more positive affect than the technology-driven baseline interaction.
- **H3:** Experience-driven teaching interactions will elicit more teaching fun than the technology-driven baseline interaction.
- **H4:** Experience-driven teaching interactions will result in increased Big Five personality traits extraversion, openness, agreeableness, and conscientiousness for the robot learner compared to the technology-driven baseline interaction (indicating a motivated learner overall, and with conscientiousness and openness particularly predicting academic success, while extraversion and agreeableness are typically predicting sociability [17, 30]).
- **H5:** Experience-driven teaching interactions will change respondents' perceived alterity relationship with the robot learner (feeling more like a counterpart than a mere tool) compared to the technology-driven baseline interaction.

5 Results

5.1 Immersion (Control)

A repeated measures ANOVA with *teaching interaction* as the factor and *immersion* as the dependent variable revealed a significant main effect of *teaching interaction*, $F(7, 875) = 5.10$, $p < .01$, $\eta_p^2 = .04$[4]. A comparison of the baseline interaction with the average of all experience-driven teaching interactions, however, showed no significant difference: an *a priori* Helmert contrast revealed a non-significant mean difference of -0.22, $SE = 0.12$, $t(875) = -1.83$, $p = .07$.

In general, experience-driven interaction did not lead to significantly more or less *immersion* ($M = 5.72$, $SD = 1.59$) compared to the baseline interaction ($M = 5.88$, $SD = 1.36$). Only two of the seven post hoc Bonferroni-corrected pairwise comparisons were significant[5]: Storytime (5): $M = 5.63$, $\delta = -0.25$, $SE = 0.08$, $p = .04$, and AsleepStudy (8): $M = 5.48$, $\delta = -0.40$, $SE = 0.09$, $p < .01$.

5.2 Need Satisfaction

A repeated measures ANOVA with *teaching interaction* as the factor and *need satisfaction* as the measure (the mean of the four

[4]Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 15.0$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.853$).

[5]For the sake of brevity, we will only report post hoc Bonferroni-corrected pairwise comparisons from here on out when they are not significant. See Appendix A.2 for a complete report of post hoc pairwise comparisons for all measures.

needs: competence, relatedness, popularity, and security) revealed a significant main effect of *teaching interaction*, $F(7, 3521) = 7.35$, $p < .01$, $\eta_p^2 < .05$[6]. A comparison of the baseline interaction with the average of all experience-driven teaching interactions was significant: an *a priori* Helmert contrast revealed a mean difference of -0.21, $SE = 0.07$, $t(3521) = -2.91$, $p < .01$. As expected (H1), experience-driven interaction led to more *need satisfaction* ($M = 5.01$, $SD = 0.95$) compared to the baseline interaction ($M = 4.80$, $SD = 0.95$), cf. Figure 3 and Table 4.

Overall, all experience-driven teaching interactions increased need satisfaction for the human teacher when compared to the purely technical teaching, Baseline (1).

5.3 Positive Affect

A repeated measures ANOVA with *teaching interaction* as the factor and *positive affect* as the measure revealed a significant main effect of *teaching interaction*, $F(7, 875) = 5.15$, $p < .001$, $\eta_p^2 < .05$[7]. A comparison of the baseline interaction with the average of all experience-driven teaching interactions was significant: an *a priori* Helmert contrast revealed a mean difference of -0.28, $SE = 0.13$, $t(875) = -2.91$, $p < .01$. As expected (H2), experience-driven interaction led to more *positive affect* ($M = 4.98$, $SD = 1.73$) compared to the baseline interaction ($M = 4.70$, $SD = 1.77$), cf. Table 4.

Experience-driven teaching interactions increased positive affect, making humans feel more active, determined, attentive, inspired, and/or alert when teaching objects to the robot. However, the three teaching interactions ProActive (2) ($M = 4.96$, $SE = 0.13$, $p = .06$), PatternML (4) ($M = 4.94$, $SE = 0.13$, $p = .06$) and AsleepStudy (8) ($M = 4.90$, $SE = 0.13$, $p = .08$) failed to differ significantly from the technology-driven teaching interaction ($M = 4.70$, $SE = 0.13$).

5.4 Teaching Fun

A repeated measures ANOVA with *teaching interaction* as the factor and *teaching fun* as the measure revealed a significant main effect of teaching interaction, $F(7, 875) = 5.98$, $p < .001$, $\eta_p^2 = 0.04$[8]. A comparison of the baseline interaction with the average of all experience-driven teaching interactions was significant: An *a priori* Helmert contrast revealed a mean difference of -0.34, $SE = 0.15$, $t(875) = -2.28$, $p < .05$. As expected (H3), experience-driven interaction led to more *teaching fun* ($M = 4.98$, $SD = 2.02$) compared to the baseline interaction ($M = 4.64$, $SD = 2.18$).

This suggests that most types of experiential teaching are effective in increasing enjoyment and engagement during the object teaching process, especially when the robot learner visualizes an object's story or recognizes its role in the teacher's routines. Only two teaching interactions did not differ significantly from the baseline in terms of *Teaching fun*: ProActive (2): $M = 4.87$, $\delta = 0.23$, $SE = 0.15$, $p = .31$; and PatternML (4): $M = 4.84$, $\delta = 0.20$, $SE = 0.15$, $p = 1.00$. Thus, designs where the robot simply initiates the teaching process or presents similarly textured objects appear less enjoyable, possibly due to their mechanical nature and lack of creative interaction.

5.5 Perceived Personality Traits of the Robot Learner

5.5.1 Extraversion. A repeated measures ANOVA with *teaching interaction* as the factor and perceived *extraversion* of the robot as the measure revealed a significant main effect of teaching interaction, $F(7, 875) = 22.20$, $p < .01$, $\eta^2 = 0.06$[9]. A comparison of the baseline interaction with the average of all experience-driven teaching interactions was significant: an *a priori* Helmert contrast revealed a mean difference of -0.74, $SE = 0.12$, $t(875) = -5.98$, $p < .01$. As expected (H4), experience-driven interaction led to increased perceived robot learner *Extraversion* ($M = 5.50$, $SD = 1.67$) compared to the baseline interaction ($M = 4.76$, $SD = 1.77$), cf. Table 4.

Participants reported that the robot was more forthcoming and communicative, thus more engaging, across all seven experience-driven teaching interactions.

5.5.2 Openness. A repeated measures ANOVA with *teaching interaction* as the factor and perceived *openness* as the measure revealed a significant main effect of teaching interaction, $F(7, 875) = 11.9$, $p < .001$, $\eta_p^2 = 0.02$[10]. A comparison of the baseline interaction with the average of all experience-driven teaching interactions was significant: an *a priori* Helmert contrast revealed a mean difference of -0.52, $SE = 0.11$, $t(875) = -4.86$, $p < .01$. As expected (H4), experience-driven interaction led to increased perceived *openness* ($M = 5.94$, $SD = 1.46$) compared to the baseline interaction ($M = 5.42$, $SD = 1.63$), cf. Table 4.

Participants reported that the robot was more open to learning and receptive to new information during the experience-driven teaching interactions.

5.5.3 Agreeableness. A repeated measures ANOVA with *teaching interaction* as the factor and perceived *agreeableness* as the measure revealed a significant main effect of teaching interaction, $F(6.27, 875) = 12.7$, $p < .001$, $\eta_p^2 = 0.06$[11]. A comparison of the baseline interaction with the average of all experience-driven teaching interactions was significant: an *a priori* Helmert contrast revealed a mean difference of -0.51, $SE = 0.11$, $t(875) = -4.67$, $p < .01$. As expected (H4), experience-driven interaction led to increased perceived *agreeableness* ($M = 5.63$, $SD = 1.46$) compared to the baseline ($M = 5.12$, $SD = 1.63$).

Participants reported the robot as being more cooperative and pleasant. Effect sizes varied, with ColorcodeML (3) and RecogHabits (6) showing the largest enhancements over baseline (cf. Table 4).

5.5.4 Conscientiousness. A repeated measures ANOVA with *teaching interaction* as the factor and *conscientiousness* as the measure

[6] Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 23.7$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.943$).
[7] Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 22.6$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.879$).
[8] Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 22.8$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.902$).
[9] Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 78.7$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.812$).
[10] Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 33.9$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.820$).
[11] Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 39.4$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.895$).

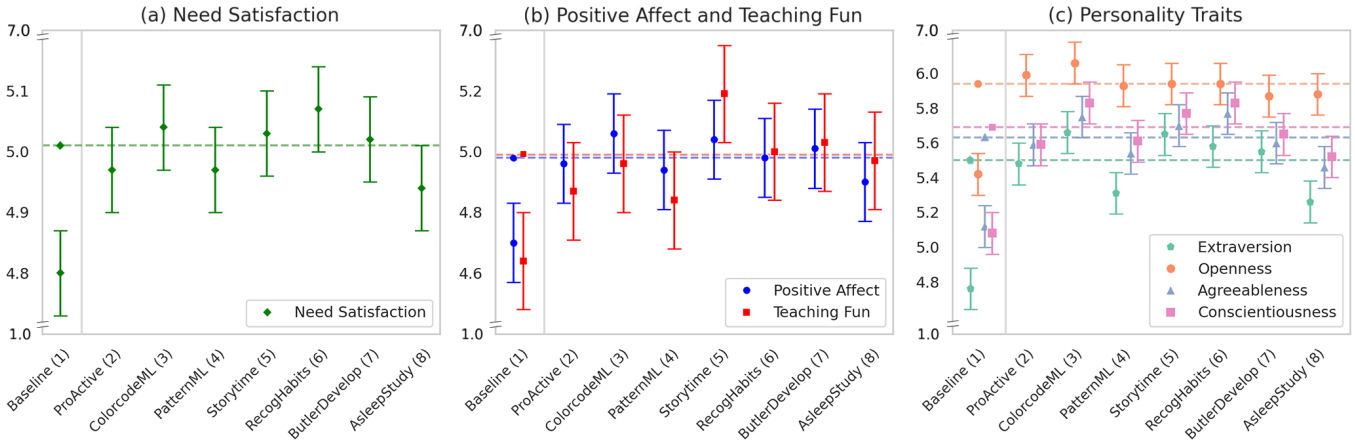


Figure 3: Mean values and variability for (a) Need Satisfaction, (b) Positive Affect and Teaching Fun, and (c) perceived robot learner personality traits (Extraversion, Openness, Agreeableness, and Conscientiousness) across all eight teaching interactions. The dashed lines represent the average mean values of the seven experience-driven interactions (2) to (8) and were used to calculate weighted Helmert contrasts, comparing the seven new teaching interactions against the one technical baseline (1).

Table 4: Average means and standard deviations for each teaching interaction measure. The column "Mean (2-8)" represents the mean for the seven experience-driven teaching interactions. A significant difference in this column indicates a significant Helmert contrast between the seven new interactions relative to the baseline interaction.

Measure	Base -line (1)	Mean (2-8)	Pro Active (2)	Colorcode ML (3)	Pattern ML (4)	Story -time (5)	Recog Habits (6)	Butler Develop (7)	Asleep Study (8)
Immersion	5.88 (1.36)	5.72 (1.59)	5.76 (1.63)	5.79 (1.63)	5.86 (1.50)	[a]5.63 (1.63)	5.80 (1.50)	5.74 (1.50)	[a]5.48 (1.77)
Need Satisfaction	4.80 (0.95)	[a]5.01 (0.95)	[a]4.97 (0.95)	[a]5.04 (0.95)	[a]4.97 (0.95)	[a]5.03 (0.95)	[a]5.07 (0.95)	[a]5.02 (0.95)	[a]4.94 (0.95)
Positive Affect	4.70 (1.77)	[a]4.98 (1.73)	4.96 (1.77)	[a]5.06 (1.63)	4.94 (1.77)	[a]5.04 (1.63)	[a]4.98 (1.77)	[a]5.01 (1.77)	4.90 (1.77)
Teaching Fun	4.64 (2.18)	[a]4.98 (2.02)	4.87 (2.04)	[a]4.96 (2.04)	4.84 (2.04)	[a]5.19 (2.04)	[a]5.00 (2.04)	[a]5.03 (1.9)	[a]4.97 (2.04)
Extraversion	4.76 (1.77)	[a]5.50 (1.67)	[a]5.48 (1.63)	[a]5.66 (1.50)	[a]5.31 (1.77)	[a]5.65 (1.63)	[a]5.58 (1.77)	[a]5.55 (1.63)	[a]5.26 (1.77)
Openness	5.42 (1.63)	[a]5.94 (1.46)	[a]5.99 (1.50)	[a]6.06 (1.36)	[a]5.93 (1.50)	[a]5.94 (1.50)	[a]5.94 (1.36)	[a]5.87 (1.50)	[a]5.88 (1.50)
Agreeableness	5.12 (1.63)	[a]5.63 (1.46)	[a]5.59 (1.50)	[a]5.75 (1.36)	[a]5.54 (1.50)	[a]5.70 (1.50)	[a]5.77 (1.36)	[a]5.60 (1.50)	[a]5.46 (1.50)
Conscientiousness	5.08 (1.63)	[a]5.69 (1.50)	[a]5.59 (1.50)	[a]5.83 (1.36)	[a]5.61 (1.50)	[a]5.77 (1.50)	[a]5.83 (1.50)	[a]5.65 (1.50)	[a]5.52 (1.63)
Alterity	5.08 (1.63)	[a]5.69 (1.50)	[a]5.59 (1.50)	[a]5.83 (1.36)	[a]5.61 (1.50)	[a]5.77 (1.50)	[a]5.83 (1.50)	[a]5.65 (1.50)	[a]5.52 (1.63)

[a] *significantly different from Baseline (1) at $p < 0.05$ or less*

revealed a significant main effect of teaching interaction, $F(7, 875) = 15.2$, $p < .001$, $\eta^2 = 0.11$. A comparison of the baseline interaction with the average of all experience-driven teaching interactions was significant: an *a priori* Helmert contrast revealed a mean difference of -0.61, $SE = 0.11$, $t(875) = -5.44$, $p < .01$[12]. As expected (H4),

[12]Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 52.1$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.848$).

experience-driven interaction led to increased perceived *conscientiousness* ($M = 5.69, SD = 1.50$) compared to the baseline interaction ($M = 5.08$, $SD = 1.63$), cf. Table 4.

Participants perceived the robot as more diligent and reliable across all experience-driven teaching interactions.

5.6 Human-Robot Technology Relationship – Alterity

A repeated measures ANOVA with *teaching interaction* as the factor and the technology relationship type *alterity* as the measure revealed a significant main effect of teaching interaction, $F(7, 875) = 8.67$, $p < .01$, $\eta^2 = 0.06$. A comparison of the baseline interaction with the average of all experience-driven teaching interactions was significant: an *a priori* Helmert contrast revealed a mean difference of -0.38, $SE = 0.10$, $t(875) = -3.83$, $p < .01$[13]. As expected (H5), experience-driven interaction led to increased *Alterity* ($M = 5.69$, $SD = 1.50$) compared to the baseline interaction ($M = 5.08$, $SD = 1.63$).

Compared to the technology-driven teaching interaction, experience-driven teaching interactions significantly intensified the perception of the robot as a counterpart—i.e., a social entity being collaborated with rather than an object or tool. Participants reported a stronger sense of exchange with the robot learner in the experiential teaching methods.

5.7 Summary

All seven experience-driven teaching methods significantly improved psychological need satisfaction, with RecogHabits (6) having the strongest effect, making participants feel more competent, connected with, and valued by their robot learner. Positive affect was significantly higher in most experience-driven interactions, with ColorcodeML (3), Storytime (5), and RecogHabits (6) particularly increasing feelings of being active, engaged, and determined during the object teaching. Most experience-driven designs also increased the fun of teaching the robot, especially Storytime (5), RecogHabits (6), and ButlerDevelop (7). However, ProActive (2) and PatternML (4) did not show a significant improvement in terms of fun certainly due to emphasizing the technical and pragmatic aspects of the robot. In addition, experience-driven interactions increased perceptions of the robot as extroverted, open, agreeable, and conscientious. This is the basis for perceiving it as more sociable, eager to learn, and diligent as a learner. This was particularly evident in the ColorcodeML (3) and Storytime (5) conditions. Overall, the results showed that teaching interactions designed based on positive practices drawn from anecdotal teaching experiences can significantly improve user engagement, enjoyment, and perceptions of the robot's personality, i.e., making the experience of teaching objects feel less tedious and more positive.

6 Discussion

We designed seven experience-driven teaching interactions for an existing object-learning robot intended for domestic use (i.e., *Johnny*) and evaluated them in a video vignette study. The general question was whether the experience-driven teaching created a more meaningful and motivating teaching experience for human teachers compared to the standard, technology-driven approach. For all experience-driven designs, we extended *Johnny*'s technology-driven object-learning procedure by incorporating emotionally resonant and positive teaching practices, inspired by interviews with teachers and coaches.

We found that all experience-driven teaching interactions (individually as well as on average) resulted in higher need satisfaction when compared to the technology-driven baseline interaction. Similarly, positive affect and teaching fun were significantly higher for almost all experience-driven interactions. This suggests that it is indeed possible to create object teaching for a robot that users find more enjoyable and motivating.

We also found higher scores for the perceived *Big Five* personality traits extraversion, openness, agreeableness, and conscientiousness of the robot, across all vignettes. Hakimi and Lesar reported that four of the *Big Five* personality traits (extraversion, openness, agreeableness, and conscientiousness) were strongly linked to academic success and capability in human learners [17, 30]. In the present study, sociable and positive personality traits were positively correlated with teaching fun (see Table 3), the most direct measure for teaching enjoyment and motivation.

6.1 Design Implications

The way humans prefer to teach is currently at odds with the technical implementation of how robots are taught. Our findings offer several ideas for interaction designers to address the artificiality and tediousness of technology-driven teaching procedures. This appears particularly important since many future robots will be highly dependent on human(-guided) input. This calls for an experience-driven design approach, where the teachers' perceptions and feelings are central to the design of the teaching interaction.

We collected anecdotes from various instructors in various teaching contexts to better understand positive teaching experiences and to serve as a blueprint. While only anecdotes, these are based on many hours of real interactions with various learners. Other than in a predominantly problem-driven approach, we urge designers to focus on successful examples of relevant interactions to better understand how to make the teacher-learner exchange more enjoyable and worthwhile (see wellbeing-driven design [25, 26]) rather than on the problem alone (i.e., tedious interaction). The present study showed that teaching interactions derived through this approach are in general more positive and engaging, which hints at their validity despite being anecdotal. Consequently, designers can base further interaction designs on these anecdotes or even collect additional anecdotes to find more ways to redesign current teaching interactions.

While all experience-driven teaching interactions were more engaging compared to the technology-driven interaction, they nevertheless highlighted different aspects. For instance, in one concept (ColorcodeML (3)), we let the robot actively share its peculiarly 'machinic' perspective on the object learned by stating its color in hexadecimal code and having it ask how the human teacher sees the object. Through this, the teacher perceived the robot as

[13] Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(7) = 19.1$, $p < .001$; therefore, we applied the Greenhouse–Geisser correction to the degrees of freedom ($\varepsilon = 0.875$).

an extroverted and curious learner, unafraid to make follow-up inquiries to learn even more. In another design (Storytime (5)), the robot asked the teacher for a personal story related to the object taught. The robot subsequently 'artistically' generated a drawing of the teacher's story. This design resulted in the highest score for teaching fun among all tested interactions, presumably because the teacher receives a personalized 'gift' in return for their teaching effort. Letting the robot not only recognize objects but also the related human routines and habits (RecogHabits (6)) resulted in particularly high psychological need satisfaction and higher perceptions of the robot's agreeableness and conscientiousness. All these designs have in common that they establish forms of reciprocal exchange and mutual understanding, yet different designs will highlight different aspects.

Interestingly, simply changing the technology-driven baseline interaction by having the robot rather than the human initiate the object teaching (ProActive (2)) did not lead to increased teaching fun. Merely demonstrating initiative in the otherwise unchanged technology-driven interaction proved less effective in improving teacher enjoyment. Thus, we recommend to design teaching interactions with a robot in a way to allow humans to better empathize and to become a part of the robot's learning journey.

In terms of practical applications, it is noteworthy that the *Johnny* robot is part of a series of robots designed as assistive technology for older adults and people with disabilities to be used in their homes. Replacing the robot's current technology-driven teaching interaction with one (or a combination) of the experience-driven interactions will certainly reduce frustration for non-expert users, will make training/personalization more appealing, will render the robot more useful, and will foster a sense of companionship—all of which could improve the quality of life for users and somewhat reduce the burden on caregivers.

In addition, such teachable robots could also take on a hybrid role in family homes, combining a household assistant with a trainable pet. They would provide both practical assistance and entertainment, especially for younger members of the household. Furthermore, the versatility of these robots could extend to service industries, such as food serving, where robots can be meaningfully trained to perform specific tasks. While each of these environments presents unique challenges and requirements, the need and potential for robots to adapt to diverse individuals and contexts remains crucial and an experience-driven teaching will support this.

6.2 Limitations

The present work has limitations. First, relying on video vignettes that present idealized scenarios with the robot's cardboard stand-in instead of real-world interactions may affect the external validity of our findings, since participants' experiences might differ with a physical robot in a real setting. While this certainly is true, vignette studies nevertheless remain a high-quality and necessary first step to better understand the effects of designed interactions. Most existing studies rarely validate interaction, fail to systematically compare different designs, and do not feature a control group or large samples.

Second, immediate post-vignette measures may not fully capture the evolution of user engagement and motivation over time. Future studies may adopt a longitudinal approach to explore how users' teaching engagement and motivation develop as they interact with assistive/service home robots over extended periods. To capture user motivations and interactions throughout this time, methods such as think-aloud protocols, experience sampling, reflective diaries, and ethnographic approaches could be used. In line with this, a recent study by Ayub and colleagues provided first insights into repeated and longitudinal interactions with continuously learning robots [7], which highlighted variations in human teaching styles and the necessity for robots to adapt accordingly. Another interesting question would be whether our experience-driven interactions can foster trust and emotional bonding, encouraging the long-term use of a robot.

Third, the study focused on the interaction in a kitchen setting, and it is unclear how experience-driven teaching might translate to other settings or group-based interactions. For instance, imagine a household where different family members teach the robot their favorite objects. Exploring how this affects the robot's recognition of certain objects, as well as the teachers' motivation and their relationships with both the robot and each other, seems certainly worth investigating. Moreover, future research could explore multi-user interactions and multi-agent applications beyond robotics, such as AI assistants that leverage the knowledge collected by the robot learner. For example, the AI assistant could use this information to personalize its responses and adapt to the specific preferences of different users.

Finally, while our sample size was sufficient for preliminary insights, it may not be representative of the broader population. Expanding the diversity of environments as well as samples in future studies will further strengthen the generalizability of our findings.

6.3 Future Work

This work shows that creating experience-driven teaching interactions for an object learning robot can result in a more positive perception of the teaching itself and the robot. A realization of the presented interaction concepts on the fully operational *Johnny* will yield further substantive insights. It would also be interesting to see how the vignette study performs compared to an evaluation of the fully implemented robot interactions—this could indicate which facets of the user experience the vignette study can capture and which ones should be tested with a real implementation. For instance, while we aimed to remedy the baseline interaction's tediousness, we must consider possible trade-offs that may emerge as a consequence of the added elements in our designs. The additional complexity of the experience-driven interaction may increase the cognitive load and/or distract from the actual object teaching task.

To further study if our design approach can be generalized and applied to other teachable robots, one could explore the tailoring of the identified teaching practices to other robot teaching methods, interactions, situations, and contexts—for example, to *programming by demonstration*, where the robot imitates the teacher's movements, or kinesthetic teaching, where the teacher guides through manual manipulation. Additionally, it would be valuable to consider robots with specialized functions, such as gripper arms used in industrial manufacturing and assembly, big agricultural robots, or

search and rescue robots, as these applications may present unique challenges and opportunities for robot teaching. It seems likely that these different learning methods and contexts may require alternative stories and experiences to engage and create meaningful motivations. Furthermore, it should also be considered how the appearance and physical form of the robot influence user perception and the metrics used in our study.

6.4 Conclusion

Our online video vignette study provided preliminary insights into designing teaching interactions for a domestic object-learning robot with the aim of making teaching a robot more positive and engaging for humans. By focusing on enhancing psychological need satisfaction, increasing positive affect, and fostering desirable learner personality traits, we have taken steps to enrich human-robot teaching interactions beyond mere instrumentality. These findings highlight the importance of user experience and emotional engagement in developing teachable robots. Ideally, this approach leads to tangible benefits for the robot's object recognition capabilities: an increase in the quality and especially quantity of object data provided by users, which progressively enhances the robot's abilities to autonomously perform tasks. On a broader scale, this contributes to more sustainable, long-term use of teachable robots.

Acknowledgments

We would like to express our gratitude to Diana Löffler and Kieu Tran, who supported study preparations, and to all the study participants. We also thank the anonymous reviewers for their constructive feedback. This work was made possible by Honda Research Institute Europe GmbH, Offenbach am Main.

References

[1] Herman Aguinis and Kyle J. Bradley. 2014. Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies. *Organizational Research Methods* 17 (10 2014), 351–371. Issue 4. https://doi.org/10.1177/1094428114547952

[2] Beatrice Alenljung, Jessica Lindblom, Rebecca Andreasson, and Tom Ziemke. 2017. User experience in social human-robot interaction. *International Journal of Ambient Computing and Intelligence* 8 (2017), 12–31. Issue 2. https://doi.org/10.4018/IJACI.2017040102

[3] Cheryl S. Alexander and Henry Jay Becker. 1978. The Use of Vignettes in Survey Research. *Public Opinion Quarterly* 42 (21 1978), 93. Issue 1. https://doi.org/10.1086/268432

[4] Brenna D. Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. 2009. A survey of robot learning from demonstration. *Robotics and Autonomous Systems* 57 (2009), 469–483. Issue 5. https://doi.org/10.1016/j.robot.2008.10.024

[5] Christiane Atzmüller and Peter M. Steiner. 2010. Experimental Vignette Studies in Survey Research. *Methodology* 6 (1 2010), 128–138. Issue 3. https://doi.org/10.1027/1614-2241/a000014

[6] Katrin Auspurg and Thomas Hinz. 2014. *Factorial Survey Experiments*. SAGE Publications Inc. https://us.sagepub.com/en-us/nam/factorial-survey-experiments/book240309

[7] Ali Ayub, Zachary De Francesco, Jainish Mehta, Khaled Yaakoub Agha, Patrick Holthaus, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. 2024. A Human-Centered View of Continual Learning: Understanding Interactions, Teaching Patterns, and Perceptions of Human Users Towards a Continual Learning Robot in Repeated Interactions. *ACM Transactions on Human-Robot Interaction* (5 2024). https://doi.org/10.1145/3659110

[8] Aude Billard and Sylvain Calinon. 2007. Handbook of Robotics Chapter 59 : Robot Programming by Demonstration. *Robotics* 48 (2007), 1371–1394. Issue 21. https://doi.org/10.1109/IROS.2008.4650593

[9] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. *Qualitative HCI Research*. Springer International Publishing. https://doi.org/10.1007/978-3-031-02217-3

[10] Maya Cakmak and Andrea L. Thomaz. 2014. Eliciting good teaching from humans for machine learners. *Artificial Intelligence* 217 (2014), 198–215. https://doi.org/10.1016/j.artint.2014.08.005

[11] Sylvain Calinon and Aude G. Billard. 2007. *What is the Teachers Role in Robot Programming by Demonstration*. 441–464.

[12] Darren M. Chan and Laurel D. Riek. 2019. Object Proposal Algorithms in the Wild: Are they Generalizable to Robot Perception? *IEEE International Conference on Intelligent Robots and Systems* (2019), 2601–2607. Issue 1. https://doi.org/10.1109/IROS40897.2019.8967830

[13] Edward L. Deci and Richard M. Ryan. 2010. *Self-Determination*. John Wiley and Sons, Inc., 1–2. https://doi.org/10.1002/9780470479216.corpsy0834

[14] Nathan Delson and Harry West. 1994. Robot programming by human demonstration: The use of human variation in identifying obstacle free trajectories. *Proceedings - IEEE International Conference on Robotics and Automation* (1994), 564–571. Issue pt 1. https://doi.org/10.1109/robot.1994.351239

[15] Hermann Dülmer. 2016. The Factorial Survey: Design Selection and its Impact on Reliability and Internal Validity. *Sociological Methods and Research* 45 (2016), 304–347. Issue 2. https://doi.org/10.1177/0049124115582269

[16] Lewis R. Goldberg. 1990. An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology* 59 (1990), 1216–1229. Issue 6. https://doi.org/10.1037/0022-3514.59.6.1216

[17] Soraya Hakimi, Elaheh Hejazi, and Masoud Gholamali Lavasani. 2011. The relationships between personality traits and students' academic achievement. *Procedia - Social and Behavioral Sciences* 29 (2011), 836–845. https://doi.org/10.1016/j.sbspro.2011.11.312

[18] Jiying Han and Hongbiao Yin. 2016. Teacher motivation: Definition, research development and implications for teachers. *Cogent Education* 3 (12 2016), 1217819. Issue 1. https://doi.org/10.1080/2331186X.2016.1217819

[19] Stephan Hasler, Jennifer Kreger, and Ute Bauer-Wersing. 2018. *Interactive Incremental Online Learning of Objects Onboard of a Cooperative Autonomous Mobile Robot*. Vol. 11307 LNCS. Springer International Publishing. 279–290 pages. https://doi.org/10.1007/978-3-030-04239-4_25

[20] Marc Hassenzahl, Susanne Boll, Jan Borchers, Astrid Rosenthal von Der Pütten, and Volker Wulf. 2019. Otherware: How to best interact with autonomous systems. (2019).

[21] Marc Hassenzahl, Sarah Diefenbach, and Anja Göritz. 2010. Needs, affect, and interactive products - Facets of user experience. *Interacting with Computers* 22 (2010), 353–362. Issue 5. https://doi.org/10.1016/j.intcom.2010.04.002

[22] Marc Hassenzahl, Kai Eckoldt, Sarah Diefenbach, Matthias Laschke, Eva Lenz, and Joonhwan Kim. 2013. Designing moments of meaning and pleasure. Experience design and happiness. *International Journal of Design* 7 (2013), 21–31. Issue 3.

[23] Marc Hassenzahl, Annika Wiklund-Engblom, Anette Bengs, Susanne Hägglund, and Sarah Diefenbach. 2015. Experience-Oriented and Product-Oriented Evaluation: Psychological Need Fulfillment, Positive Affect, and Product Perception. *International Journal of Human-Computer Interaction* 31 (2015), 530–544. Issue 8. https://doi.org/10.1080/10447318.2015.1064664

[24] Don Ihde. 1990. *Technology and the Lifeworld: From Garden to Earth*. Indiana University Press (1990). https://philpapers.org/rec/IHDTAT-3

[25] Holger Klapperich, Marc Hassenzahl, Melanie Becker, Diana Cürlis, Thorsten Frackenpohl, Henning Köhler, Kai Ludwigs, and Marius Tippkämper. 2019. *Mind the Gap: A Social Practice Approach to Wellbeing-Driven Design*. Routledge, 154–169. https://doi.org/10.4324/9781315121383

[26] Holger Klapperich, Matthias Laschke, and Marc Hassenzahl. 2018. The positive practice canvas - Gathering inspiration for wellbeing-driven design. *ACM International Conference Proceeding Series* (2018), 74–81. https://doi.org/10.1145/3240167.3240209

[27] Lenneke Kuijer, Annelise De Jong, and Daan Van Eijk. 2013. Practices as a unit of design: An exploration of theoretical guidelines in a study on bathing. *ACM Transactions on Computer-Human Interaction* 20 (2013). Issue 4. https://doi.org/10.1145/2493382

[28] Matthias Laschke, Robin Neuhaus, Judith Dörrenbächer, Marc Hassenzahl, Volker Wulf, Astrid Rosenthal-Von Der Pütten, Jan Borchers, and Susanne Boll. 2020. Otherware needs Otherness: Understanding and Designing Artificial Counterparts. *ACM International Conference Proceeding Series* (2020), 6–9. https://doi.org/10.1145/3419249.3420079

[29] Hoo Man Lee and Joong Bae Kim. 2013. A survey on robot teaching: Categorization and brief review. *Applied Mechanics and Materials* 330 (2013), 648–656. Issue June. https://doi.org/10.4028/www.scientific.net/AMM.330.648

[30] Irena Lesar, Ivan Čuk, and Mojca Peček. 2014. Teachers on perceived traits and academic achievements of regular pupils and pupils with special needs in mainstream primary schools. *Teachers and Teaching: Theory and Practice* 20 (2014), 358–374. Issue 3. https://doi.org/10.1080/13540602.2013.848520

[31] Gabriel Lima, Nina Grgić-Hlača, and Meeyoung Cha. 2021. Human Perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3411764.3445260

[32] Christoph Lutz and Aurelia Tamò-Larrieux. 2021. Do Privacy Concerns About Social Robots Affect Use Intentions? Evidence From an Experimental Vignette Study. *Frontiers in Robotics and AI* 8 (4 2021). https://doi.org/10.3389/frobt.2021.627958

[33] Stanislava Naneva, Marina Sarda Gou, Thomas L. Webb, and Tony J. Prescott. 2020. A Systematic Review of Attitudes, Anxiety, Acceptance, and Trust Towards Social Robots. *International Journal of Social Robotics* 12 (12 2020), 1179–1201. Issue 6. https://doi.org/10.1007/s12369-020-00659-4

[34] Pierre Rouanet, Pierre Yves Oudeyer, Fabien Danieau, and David Filliat. 2013. The impact of human-robot interfaces on the learning of visual objects. *IEEE Transactions on Robotics* 29 (2013), 525–541. Issue 2. https://doi.org/10.1109/TRO.2012.2228134

[35] Aran Sena, Yuchen Zhao, and Matthew J. Howard. 2018. Teaching human teachers to teach robot learners. *Proceedings - IEEE International Conference on Robotics and Automation*, 5675–5681. https://doi.org/10.1109/ICRA.2018.8461194

[36] Elizabeth Shove, Mika Pantzar, and Matt Watson. 2012. *The Dynamics of Social Practice.* SAGE Publications Ltd. https://us.sagepub.com/en-us/nam/the-dynamics-of-social-practice/book235021

[37] Catherine Sinclair. 2008. Initial and changing student teacher motivation and commitment to teaching. *Asia-Pacific Journal of Teacher Education* 36 (5 2008), 79–104. Issue 2. https://doi.org/10.1080/13598660801971658

[38] Andrea L. Thomaz and Cynthia Breazeal. 2008. Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence* 172 (2008), 716–737. Issue 6-7. https://doi.org/10.1016/j.artint.2007.09.009

[39] Edmund R. Thompson. 2007. Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology* 38 (2007), 227–242. Issue 2. https://doi.org/10.1177/0022022106297301

[40] Anna Lisa Vollmer and Lars Schillingmann. 2018. On Studying Human Teaching Behavior with Robots: a Review. *Review of Philosophy and Psychology* 9 (2018), 863–903. Issue 4. https://doi.org/10.1007/s13164-017-0353-4

[41] Anna Lisa Vollmer, Britta Wrede, Katharina J. Rohlfing, and Pierre Yves Oudeyer. 2016. Pragmatic frames for teaching and learning in human-robot interaction: Review and challenges. *Frontiers in Neurorobotics* 10 (2016), 1–20. Issue OCT. https://doi.org/10.3389/fnbot.2016.00010

[42] Chao Wang, Stephan Hasler, Manuel Mühlig, Frank Joublin, and Lydia Fischer. 2020. Designing Interaction for Multi-agent System in an Office Environment. (2020). Issue February.

[43] Weitian Wang, Yi Chen, Rui Li, and Yunyi Jia. 2019. Learning and comfort in human-robot interaction: A review. *Applied Sciences (Switzerland)* 9 (2019). Issue 23. https://doi.org/10.3390/app9235152

[44] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54 (1988), 1063–1070. Issue 6. https://doi.org/10.1037/0022-3514.54.6.1063

[45] Heiko Wersing, Stephan Kirstein, Michael Götting, Holger Brandl, Mark Dunn, Inna Mikhailova, Christian Goerick, Jochen Steil, Helge Ritter, and Edgar Körner. 2007. Online learning of objects in a biologically motivated visual architecture. *International Journal of Neural Systems* 17 (2007), 219–230. Issue 4. https://doi.org/10.1142/S0129065707001081

A Appendix

A.1 Complete Questionnaire with All Items

Table A.1: Full set of questionnaire items that participants responded to after viewing each video vignette.

Category	Item	Construct	Response Scale
Demographics	Your gender	-	-
	Your age in years	-	-
Immersion Check	I was able to put myself into the situation shown in the video.	Immersion	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I could relate to the scenario presented in the video.	Immersion	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
Need Satisfaction [13, 23]	I felt that I was completing the task of teaching the robot successfully.	Competency	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt that I was taking on and mastering the challenge of teaching the robot new information.	Competency	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt a sense of contact with the robot, whom I care for.	Relatedness	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt close and connected with the robot.	Relatedness	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt that I was a person others seek out and admire.	Popularity	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt that I am someone others take as guidance.	Popularity	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt glad that my life is structured.	Security	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt glad that I have a comfortable set of routines and habits.	Security	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely

Continued on next page

Table A.1 – Continued from previous page

Category	Item	Construct	Response Scale
Positive Affect (PANAS Short Form) [39]	Teaching the robot made me feel active.	Active	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	Teaching the robot made me feel determined.	Determined	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	Teaching the robot made me feel attentive.	Attentive	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	Teaching the robot made me feel inspired.	Inspired	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	Teaching the robot made me feel alert.	Alert	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
Teaching Fun	Teaching the robot gave me pleasure.	Enjoyability of Teaching	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	Teaching the robot amused me.	Enjoyability of Teaching	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely

Continued on next page

Table A.1 – Continued from previous page

Category	Item	Construct	Response Scale
Big Five Personality Traits of Robot Learner Based on [16, 17, 30]	I felt the robot was active and took initiative.	Extraversion	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt the robot was communicative and expressive.	Extraversion	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt the robot was open to learn new things.	Openness	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt the robot was receptive for new information.	Openness	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt the robot was very cooperative and forthcoming.	Agreeableness	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt the robot was considerate of me and what I taught.	Agreeableness	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt the robot was perceptive of what was happening.	Conscientiousness	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt the robot was paying attention to detail.	Conscientiousness	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt that the robot was quirky.	Neuroticism	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt that the robot had a very peculiar way to cope with the situation.	Neuroticism	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
Technology-Relationship Alterity [24]	I felt that the robot was a discrete thing separate from me.	Alterity	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt that I engaged in a dialogue with the robot.	Alterity	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt that the robot was acting autonomously.	Alterity	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely
	I felt that I act in cooperation with the robot.	Alterity	Not at all, Hardly at all, Little, Neutral, Good, Very good, Extremely

A.2 Post-Hoc Pairwise Comparisons for All Measures

A.2.1 Immersion (Control). Post hoc Bonferroni-corrected pairwise comparisons between the baseline interaction (1) and two of the seven experience-driven teaching interactions revealed significant differences: (Storytime (5): M = 5.63, δ = -0.25, SE = 0.08, p = .04; AsleepStudy (8): M = 5.48, δ = -0.40, SE = 0.09, p < .01). However, the remaining five teaching interactions did not differ significantly from the baseline (ProActive (2): M = 5.76, δ = -0.12, SE = 0.09, p = 1.00; ColorcodeML (3): M = 5.79, δ = -0.09, SE = 0.04, p = 1.00; PatternML (4): M = 5.86, δ = -0.02, SE = 0.06, p = 1.00; RecogHabits (6): M = 5.80, δ = -0.08, SE = 0.07, p = 1.00; ButlerDevelop (7): M = 5.74, δ = -0.14, SE = 0.08, p = 1.00).

A.2.2 Need Satisfaction. Post hoc Bonferroni-corrected pairwise comparisons between the baseline interaction and the seven experience-driven teaching interactions revealed significant differences: ProActive (2): M = 4.97, δ = -0.16, SE = 0.04, t(503) = -3.91, p < 0.01; ColorcodeML (3): M = 5.04, δ = -0.23, SE = 0.041, t(503) = -5.76, p < .01; PatternML (4): M = 4.97, δ = -0.17, SE = 0.04, t(503) = -3.97, p < 0.01; Storytime (5): M = 5.03, δ = -0.22, SE = 0.05, t(503) = -4.62, p < .01; RecogHabits (6): M = 5.07, δ = -0.26, SE = 0.05, t(503) = -5.85, p < .01; ButlerDevelop (7): M = 5.02, δ = -0.21, SE = 0.05, t(503) = -4.70, p < .01; AsleepStudy (8): M = 4.94, δ = -0.14, SE = 0.05, t(503) = -3.14, p < 0.05.

A.2.3 Positive Affect. Further post hoc Bonferroni-corrected pairwise comparisons revealed significant differences between the baseline interaction and four experience-driven teaching interactions: ColorcodeML (3): M = 5.06, δ = 0.36, SE = 0.124, p < .01; Storytelling (5): M = 5.04, δ = 0.35, SE = 0.124, p < .01; RecogHabits (6): M = 4.98, δ = 0.28, SE = 0.126, p < .01; ButlerDevelop (7): M = 5.01, δ = 0.31, SE = 0.126, p < .01.

A.2.4 Teaching Fun. Further post hoc Bonferroni-corrected pairwise comparisons between the technical teaching interaction Baseline (1) and the seven experiential teaching interactions revealed significant differences for the following five teaching interactions: ColorcodeML (3): M = 4.96, δ = 0.32, SE = 0.15, p < .05; Storytime (5): M = 5.19, δ = 0.55, SE = 0.15, p < .01; RecogHabits (6): M = 5.00, δ = 0.36, SE = 0.15, p < .01; ButlerDevelop (7): M = 5.03, δ = 0.39, SE = 0.14, p < .01; AsleepStudy (8): M = 4.97, δ = 0.33, SE = 0.15, p < .05. The remaining two teaching interactions did not differ significantly in terms of Teaching Fun experienced by the human teacher: ProActive (2): M = 4.87, δ = 0.23, SE = 0.15, p = .31; PatternML (4): M = 4.84, δ = 0.20, SE = 0.15, p = 1.00.

A.2.5 Big Five Personality Traits of the Robot Learner.

Extraversion. Further post hoc Bonferroni-corrected pairwise comparisons between the baseline interaction and the seven experience-driven teaching interactions revealed significant differences: ProActive (2): M = 5.48, δ = 0.72, SE = 0.12, p < .01; ColorcodeML (3): M = 5.66, δ = 0.90, SE = 0.11, p < .01; PatternML (4): M = 5.31, δ = 0.55, SE = 0.13, p < .01; Storytime (5): M = 5.65, δ = 0.89, SE = 0.12, p < .01; RecogHabits (6): M = 5.58, δ = 0.81, SE = 0.13, p < .01; ButlerDevelop (7): M = 5.55, δ = 0.79, SE = 0.12, p < .01; AsleepStudy (8): M = 5.26, δ = 0.50, SE = 0.13, p < .01.

Openness. Further post hoc Bonferroni-corrected pairwise comparisons between the baseline interaction and the seven experience-driven teaching interactions revealed significant differences: ProActive (2): M = 5.99, δ = 0.57, SE = 0.11, p < .01; ColorcodeML (3): M = 6.06, δ = 0.64, SE = 0.10, p < .01; PatternML (4): M = 5.93, δ = 0.51, SE = 0.11, p < .01; Storytime (5): M = 5.94, δ = 0.52, SE = 0.11, p < .01; RecogHabits (6): M = 5.94, δ = 0.53, SE = 0.10, p < .01; ButlerDevelop (7): M = 5.87, δ = 0.45, SE = 0.11, p < .01; AsleepStudy (8): M = 5.88, δ = 0.46, SE = 0.11, p < .01.

Agreeableness. Further post hoc Bonferroni-corrected pairwise comparisons between the baseline interaction and the seven experience-driven teaching interactions revealed significant differences: ProActive (2): M = 5.59, δ = 0.47, SE = 0.09, p < .01; ColorcodeML (3): M = 5.75, δ = 0.64, SE = 0.09, p < .01; PatternML (4): M = 5.54, δ = 0.43, SE = 0.09, p < .01; Storytime (5): M = 5.70, δ = 0.59, SE = 0.10, p < .01; RecogHabits (6): M = 5.77, δ = 0.66, SE = 0.10, p < .01; ButlerDevelop (7): M = 5.60, δ = 0.48, SE = 0.09, p < .01; AsleepStudy (8): M = 5.46, δ = 0.35, SE = 0.09, p < 0.01.

Conscientiousness. Further post hoc Bonferroni-corrected pairwise comparisons between the baseline interaction and the seven experience-driven teaching interactions revealed significant differences: ProActive (2): M = 5.59, δ = 0.51, SE = 0.10, p < .01; ColorcodeML (3): M = 5.83, δ = 0.75, SE = 0.10, p < .01; PatternML (4): M = 5.61, δ = 0.53, SE = 0.09, p < .01; Storytime (5): M = 5.77, δ = 0.69, SE = 0.10, p < .01; RecogHabits (6): M = 5.83, δ = 0.75, SE = 0.10, p < .01; ButlerDevelop (7): M = 5.65, δ = 0.57, SE = 0.10, p < .01; AsleepStudy (8): M = 5.52, δ = 0.44, SE = 0.10, p < .01.

A.2.6 Human-Robot Technology Relationship – Alterity. Further post hoc Bonferroni-corrected pairwise comparisons between the baseline interaction and the seven experience-driven teaching interactions revealed significant differences: ProActive (2): M = 5.04, δ = 0.38, SE = 0.09, p < .01; ColorcodeML (3): M = 5.12, δ = 0.46, SE = 0.10, p < .01; PatternML (4): M = 4.99, δ = 0.33, SE = 0.10, p < .01; Storytime (5): M = 5.04, δ = 0.38, SE = 0.10, p < .01; RecogHabits (6): M = 5.12, δ = 0.46, SE = 0.10, p < .01; ButlerDevelop (7): M = 4.99, δ = 0.33, SE = 0.10, p < .01; AsleepStudy (8): M = 4.98, δ = 0.32, SE = 0.11, p < 0.01.