CHILD DEVELOPMENT

Child Development, xxxx 2020, Volume 00, Number 0, Pages 1–17

Preschoolers' Motivation to Over-Imitate Humans and Robots

Hanna Schleihauf
Max Planck Institute for Human Cognitive and Brain Sciences and University of California and German Primate Center—Leibniz Institute for Primate Research and Georg-August-University

Stefanie Hoehl
Max Planck Institute for Human Cognitive and Brain Sciences and University of Vienna

Neli Tsvetkova
New Bulgarian University

Alexander König
Ruprecht-Karls-Universität

Katja Mombaur
Ruprecht-Karls-Universität and University of Waterloo

Sabina Pauen
Ruprecht-Karls-Universität

From preschool age, humans tend to imitate causally irrelevant actions—they over-imitate. This study investigated whether children over-imitate even when they know a more efficient task solution and whether they imitate irrelevant actions equally from a human compared to a robot model. Five-to-six-year-olds (N = 107) watched either a robot or human retrieve a reward from a puzzle box. First a model demonstrated an inefficient (Trial 1), then an efficient (Trial 2), then again the inefficient strategy (Trial 3). Subsequent to each demonstration, children copied whichever strategy had been demonstrated regardless of whether the model was a human or a robot. Results indicate that over-imitation can be socially motivated, and that humanoid robots and humans are equally likely to elicit this behavior.

Imitation plays a crucial role in social interactions and provides a highly effective strategy for acquiring new skills, behaviors, and norms (e.g., Jones, 2009; Meltzoff & Marshall, 2018; Tomasello, Carpenter, Call, Behne, & Moll, 2005). In contrast to nonhuman primates, human children even copy actions that are perceivably irrelevant to attain a given goal (Clay & Tennie, 2017; Horner & Whiten, 2005). This phenomenon is called over-imitation (henceforth "oi"; Lyons, Damrosch, Lin, Macris, & Keil, 2011; Lyons, Young, & Keil, 2007) and has typically been investigated using a paradigm in which participants first watch a model retrieving a reward from a puzzle box using a combination of irrelevant actions (e.g., tapping on the surface of the box) and relevant actions (e.g., opening a lid to retrieve a reward). After observing the model demonstrating irrelevant and relevant actions, children (McGuigan & Whiten, 2009; Nielsen & Tomaselli, 2010) as well as adults (Flynn & Smith, 2012; McGuigan, Makinson, & Whiten, 2011; Whiten et al., 2016) tend to copy both, relevant and irrelevant actions, thus applying an inefficient strategy. Crucially, we only speak of oi, if irrelevant actions are performed more frequently following a social demonstration than spontaneously (i.e., in a baseline condition without prior demonstration; Hoehl, Zettersten, Schleihauf, Grätz, & Pauen, 2014; Lyons et al., 2007, 2011).

Several explanations for the phenomenon of oi are discussed in the literature (for an overview see Hoehl et al., 2019). While some accounts suggest that children over-imitate because they erroneously interpret all intentional actions as being causally relevant

We thank all the kindergartens that supported our work and all the children that participated in the study. Furthermore, we thank Shameem Wagner for proofreading the manuscript and the editing support. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG; grant number HO 4342/8-1) and a stipend from the Studienstiftung des deutschen Volkes awarded to Hanna Schleihauf.

Conflict of interest: The authors have no competing interests to declare.

Open access funding enabled and organized by Projekt DEAL.

[Corrections added on August 08, 2020, after first online publication. the open access funding statement has been added.]

Correspondence concerning this article should be addressed to Hanna Schleihauf, University of California, Berkeley, CA 94704. Electronic mail may be sent to schleihauf@berkeley.edu.

0009-3920/2020/xxxx-xxxx
DOI: 10.1111/cdev.13403

(Lyons et al., 2011), others emphasize potential social motivations: Children may want to affiliate with another person (Nielsen & Blank, 2011), or they may assume that the demonstrated actions are normative (Kenward, 2012; Kenward, Karlsson, & Persson, 2011; Keupp, Behne, & Rakoczy, 2013). Some integrative accounts suggest that children's imitation strategies are flexible; they can depend on the goal that dominates in a given situation (Over & Carpenter, 2012) or on previous experiences (Williamson & Meltzoff, 2011; Williamson, Meltzoff, & Markman, 2008). Children may follow an instrumental goal based on the motivation to reach the reward, or they may follow a social goal based on the motivation to, for example, please the experimenter or to create a social bond with the experimenter. They may also follow a combination of social and instrumental goals. It is assumed that the activation of a social goal enhances children's oi and increases their sensitivity to cultural information, such as rituals, conventions or norms (Over & Carpenter, 2012). Interestingly, in some circumstances children also tend to copy precise means when they had initial instrumental goals. Children who first manipulated an object on their own and thereby had a more difficult prior experience were more likely to imitate the precise means of a subsequently demonstrated action than children with an easier prior experience (Williamson & Meltzoff, 2011; Williamson et al., 2008).

From an evolutionary perspective, oi is often considered crucial for explaining the emergence of cultural learning and ritual transmission unique to humans (Legare, 2017; Nielsen & Tomaselli, 2010; Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009). Research along these lines found that children learn culturally relevant behaviors by selectively copying reliable and presumably knowledgeable individuals. Different characteristics, such as status, prestige, success, age, and familiarity of the model, are known to influence children's tendency to imitate irrelevant actions (e.g., McGuigan, 2013; Wood, Kendal, & Flynn, 2012; Wood et al., 2016). If the function of oi was the transmission of culturally relevant knowledge, it would seem reasonable to copy only members of one's own species because only they potentially possess culturally relevant knowledge. This is not necessarily true for other kinds of interactive partners, such as robots.

To better understand the nature of oi it thus seems important to clarify under which conditions this phenomenon can be elicited. The first goal of this study was to disentangle the contribution of instrumental and social goals in a given demonstration-imitation task. To address this issue, we adapted an existing oi paradigm to manipulate knowledge about the efficient strategy using a within-subject design. The second goal was to specify model characteristics supporting oi. Therefore, we compared oi rates following the demonstration by a human experimenter and a humanoid robot using a between-subject design.

Is oi Driven by Social or Instrumental Goals?

When children observe an inefficient strategy including irrelevant and relevant action steps, but no alternative more efficient solution, it remains unclear why they over-imitate. It could be that they follow an instrumental goal, but simply lack the causal knowledge needed to realize that some actions are irrelevant. Alternatively, they may pursue a social goal, such as affiliating with the experimenter or following an assumed social norm. To discriminate between instrumental and social goals, some researchers demonstrated not only an inefficient way to solve a task, but also the most efficient one. Since children were informed about the efficient way to solve a problem at hand, it could be ruled out that oi resulted from the fact that children misinterpreted irrelevant actions as being causally relevant (Hoehl et al., 2014; Nielsen & Blank, 2011; Nielsen, Kapitány, & Elkins, 2015; Nielsen, Tomaselli, & Kapitány, 2018; Schleihauf, Graetz, Pauen, & Hoehl, 2018). In a study by Nielsen and Blank (2011), children observed two experimenters, one using an efficient strategy and the other using an inefficient strategy to extract a reward from a box. Children only over-imitated when the inefficient experimenter later handed the box to them. The authors concluded that children have a strong affiliation (social) motivation that guides their behavior and explains the occurrence of oi. Somewhat contradicting this view, other studies revealed that children preferred the efficient strategy when they first learned the efficient way to solve a problem, even when a subsequently inefficient strategy was socially highlighted (Schleihauf et al., 2018). This points to the relevance of both instrumental and social goals in oi settings. It also reveals that their specific interplay is not yet fully understood.

One potential way to shed further light on this interplay is to sequentially present three test trials. In each trial the child observed a communicative model demonstrating one way to attain a reward from a puzzle box before the child had a go. In Trial 1, the model demonstrated a combination of irrelevant and relevant actions (i.e., using an inefficient strategy). In Trial 2, the model demonstrated only actions needed to solve the problem at hand (i.e., using the efficient strategy). In Trial 3, the model again demonstrated the inefficient strategy. The third trial was of particular interest in the present context because it allowed us to conduct a within-subject comparison of oi performance between Trial 1 (i.e., without previous demonstration of the efficient strategy) and Trial 3 (i.e., following the demonstration of the efficient strategy in Test trial 2).

Do Children (Over-)imitate Robots?

In modern times, we not only interact with humans, but also increasingly with machines, computers, and robots. More and more children are growing up with technical devices that allow for dynamic exchanges of information. Humanoid robots are being used in educational settings, as interactive guides at science museums (Shiomi, Kanda, Ishiguro, & Hagita, 2006), or as peer tutors at schools in Japan (Kanda, Hirano, Eaton, & Ishiguro, 2004) and the United States (Meltzoff, Kuhl, Movellan, & Sejnowski, 2009; Tanaka, Cicourel, & Movellan, 2007). There is also a growing body of evidence showing the benefits of therapy with robots for children with autism and Down syndrome (Warren et al., 2015). This seems rather interesting because the categorical status of robots is not entirely clear: On the one hand, they are man-made, typically have a plastic or metal surface, and usually have a certain function. On the other hand, they often show a similar body configuration to humans, move on their own, communicate, and interact and complete tasks in similar ways as people do.

To our knowledge, only one published study examined whether young children imitate robots. Itakura et al. (2008) examined whether toddlers (24–35 months old) would imitate a robot's goal-directed action that was demonstrated in a failed attempt paradigm. The results were positive. Interestingly, children were more likely to imitate a robot showing human-like congruent gaze behavior than one not showing this behavior. Even though the demonstration was on video and the robot did not display any ostensive cues (i.e., contacting the child, smiling and talking to her), children were able to infer the robot's "action goals" and were motivated to complete the demonstrated action. If the model performed the action successfully, congruent gaze was not necessary to induce children's imitation. Based on these findings, one may speculate that children will also show oi following the demonstration of a communicative humanoid robot that establishes eye contact before starting the action sequence.

Only one previous study reported oi with another nonhuman model (i.e., a puppet). Although oi rates were higher when a human rather than a puppet demonstrated irrelevant actions, children also copied the behavior of the puppet (McGuigan & Robertson, 2015). Importantly, the child could see that the puppet was acting under the control of a human experimenter. Differing from this situation, robots appear to act on their own without visible external control. Hence, it is well possible that they are perceived differently. The animate–inanimate distinction is highly prominent in children's thinking and develops early in life (see Legerstee, 1992 for a review; Rakison & Poulin-Dubois, 2001). By 5 years of age children can discriminate animate beings from inanimate objects based on multiple features, including their appearance and behavior (e.g., Pauen, 2000). A human model controlling a puppet may thus be more likely to elicit a social goal (leading to higher oi) than a robot model showing self-initiated actions, provided that the robot is categorized as a technical device. In line with this, Sommer et al. (2019) found that children showed less moral concern when a robot was mistreated than when a living being was mistreated, but afforded more moral concern to robots than to nonliving entities.

However, it is also possible that oi does not depend on the presence of a human model, but instead on the presence of certain social cues triggering social goals to become engaged in observational learning and imitation. In this study, we thus used a humanoid robot that moved and interacted similarly to an animate social being. It acted ostensively, "looking" toward the camera and apparently addressing the child. These cues are thought to be strong indicators for the transmission of cultural knowledge (Csibra & Gergely, 2011), even in the absence of a real human agent (Deligianni, Senju, Gergely, & Csibra, 2011). In line with this, infants and children are more attentive and receptive to robots displaying nonverbal contingency and social referencing. Preschoolers (but not infants, O'Connell, Poulin-Dubois, Demke, & Guay, 2009) even seem to treat them as social informants (Breazeal et al., 2016; Meltzoff, Brooks, Shon, & Rao, 2010). Furthermore, humanoid robots were shown to have a positive impact on older children's performance in cognitive tasks, but only if they showed human-like communicative behavior (André et al., 2014). All these different findings suggest that the quality of interaction between children and robots is modulated by the presence of human-like characteristics of the robot's appearance (having two arms, legs, and eyes) and also its behavior (self-initiated, more or less dynamic motion, contingent responding and ostensive cues like congruent eye gaze).

Hypotheses

This study focuses on two important research questions regarding the role of social goals for oi:

1. Do 5- to 6-year-olds over-imitate based on a social goal, and will they therefore flexibly switch back to an inefficient strategy even after having performed a more efficient strategy? We reasoned that if instrumental goals dominated children's actions oi should decrease from Trial 1 to Trial 3, because they already knew the efficient strategy after participating in Trial 2. If social goals dominated children's action, we expected to find similar oi performance in Trial 1 and Trial 3. Social goals should lead to faithful imitation of either strategy (despite having the knowledge of a more efficient solution) and to a flexible adaptation of children's responses across trials. This hypothesis is based on previous findings, in which children's tendency to over-imitate after a subsequent efficient demonstration varied with the communicative behavior of the second demonstrator. If the demonstrator who demonstrated the efficient strategy was communicative, children switched to the efficient strategy; when the efficient strategy demonstrator was not communicative, children continued to over-imitate (Hoehl et al., 2014; Schleihauf et al., 2018). Thus, in a social context, we expect children to flexibly adopt the demonstrated strategy.
2. Are children in this age range equally willing to learn an inefficient strategy from a humanoid robot as they are from a human experimenter? In Table 1 we present four possible outcomes of the study and respective conclusions. Based on the arguments raised so far, we make the following predictions: For Test trial 1 (inefficient demonstration) in the human condition, we expect a higher rate of imitation of irrelevant actions than in a baseline condition without previous demonstration, thus replicating the findings of previous studies using video demonstrations (McGuigan et al., 2011; Wood et al., 2012). With regard to the robot condition, we have no clear expectations, but speculated that oi rates will be similar to the human condition because we use a humanoid communicative robot and because the children have not a seen a more efficient solution before. For Trial 2 (efficient demonstration), we expect children to switch to the more efficient strategy, regardless of the condition. This would be consistent with Hoehl et al. (2014), who found that children switched from an inefficient to an efficient strategy if the latter was demonstrated ostensively. Outcomes for Trial 3 (inefficient demonstration) are most informative with regard to our research questions. Assuming that oi is mostly socially motivated, it still needs to be answered how that motivation is triggered. To what extent children prioritize their social goals over purely instrumental goals is likely to be depend on their perception of the robot or human demonstrating the action. Here, we see two potential outcomes: If socially motivated oi is restricted to the transmission of action knowledge between humans, then it should only occur in the human condition. However, if socially motivated oi is dependent on social cues, such as a humanoid appearance, gaze, and speech, then oi rates should be similar in both the human condition and the robot condition.

Method

Design

We used a mixed design, with between-subject (conditions: human, robot) as well as within-subject (Test trials 1, 2, and 3) factors (see Figure 1).

Participants

The data of N = 107 children aged from 5.00 to 6.65 years (M = 5.59 years) were included in the final sample of this study. By the age of 5, oi is a robust phenomenon (McGuigan, Whiten, Flynn, & Horner, 2007), thus, we tested children in this age group. Previous work with the same kind of puzzle box also tested 5- to 6-year-olds. To allow for qualitative comparisons of findings across studies, we decided to test children of the same age range. The final sample contained 29 (17 boys) children in the human condition, 34 children (14 boys) in the robot condition and 44 children in the baseline condition (21 boys). Data were collected between January 2017 and May 2019. All subjects spoke German as their first or second language and came from mixed socioeconomic backgrounds. Children from seven different kindergartens in a medium-sized German

Table 1
Possible Outcomes and Respective Interpretations

	Possible outcomes		
	Test trial 1	Test trial 3	Interpretations
Human cond.	oi	oi	oi is socially motivated and only occurs within the human species. No social motivation is elicited by a robot model.
Robot cond.	no oi	no oi	
Human cond.	oi	oi	oi is due to causal erroneous reasoning in Test trial 1, but socially motivated in Test trial 3. No social motivations are elicited by a robot model. Transmission of socially relevant behaviors only within the human species
Robot cond.	oi	no oi	
Human cond.	oi	no oi	oi is based on erroneous causal reasoning. No effect of social motivations on oi after the irrelevancy of actions is revealed
Robot cond.	oi	no oi	
Human cond.	oi	oi	Unclear whether oi in Test trial 1 is due to erroneous causal reasoning or social motivation. oi in Test trial 3 is socially motivated. Social motivations can be elicited by a humanoid robot model. Transmission of socially relevant behaviors not only within the human species as long as ostensive cues are provided
Robot cond.	oi	oi	

town were tested with the consent of their parents. Five additional children were recruited from a database of children who had previously participated in other psychological studies. The data from additional 25 children were not included in the analysis—one due to experimenter error (robot condition) and eight because the child could not complete the task of extracting a golden marble from the box in Test trial 1 of the experiment (n = 4 in the human and n = 4 in the robot condition) or in the baseline condition (n = 16). This exclusion criterium was applied to ensure that the experience of failure did not have any effect on imitation rates. Overall, the children and their families had a mixed socioeconomic background.

Material

The Puzzle Box

Children were presented with a transparent puzzle box containing golden marbles. All marbles were placed inside a clear horizontal tube, and each one had a small magnet attached to it so that it could be extracted from the box using a magnetic rod. There was a black wooden lever on the top of the box and a black button attached to the side of the box (see Figure 2; Hoehl et al., 2014; Schleihauf et al., 2018). Due to the transparency of the box, it was possible to observe that any manipulation of both elements did not affect the tube or the marble inside.

The Robot

The robot used in the videos is a humanoid robot (type Nao Evolution [Figure 2]) and was produced in 2011 by Aldebaran Robotics. It is 58 cm tall and has 25 *df*. Important features include its LED light-enhanced eyes that can change color to create the illusion of congruent eye contact and goal-directed gaze.

The Videos

We decided to use videos rather than live presentations to ensure standardization of demonstrations across conditions and to avoid dropouts due to malfunctioning of the robot. A fairly large body of research indicates that children by the age of five are equally willing to copy demonstrated actions from a live-demonstration as from a video demonstration (McGuigan et al., 2011; Wood et al., 2012). Videos showed either the robot or a woman who was unfamiliar and different from the one who ran the experiment. Each video included multiple takes to show different parts of the action as close-ups. To obtain comparability between the two experimental conditions, the number and duration of different takes as well as the positions of the video cameras were identical in both the robot and human model conditions.

Three different scenes were videotaped per condition: In the video for Trial 1 (56 s) the model greets the child, introduces herself, and explains

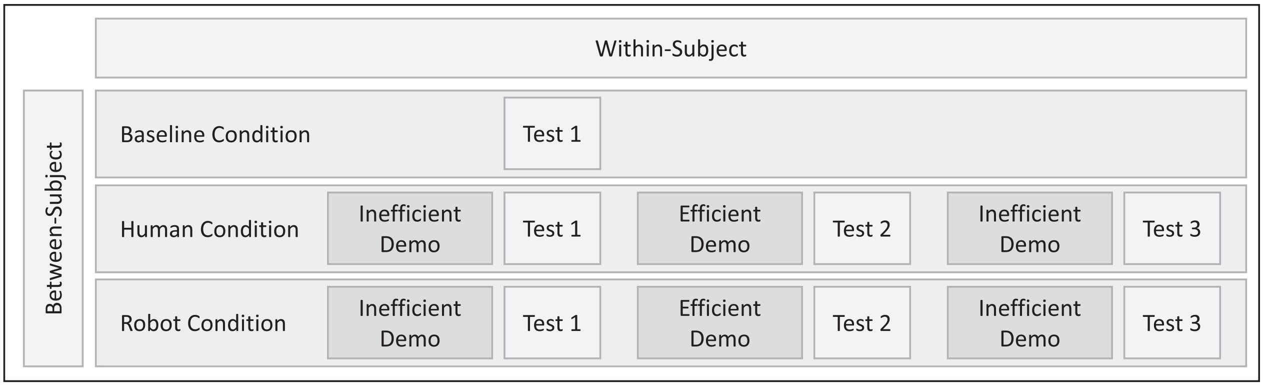

Figure 1. Conditions.

Figure 2. Actions performed on the puzzle box in both the human condition and robot condition. A: irrelevant action - moving a non-functional lever from one side to the other; B: irrelevant action - pushing a nonfunctional button; C: relevant action - opening a flap and inserting a magnetic rod to extract a marble.

that there are golden marbles inside the box which can be exchanged for stickers. The human actor speaks with her normal voice and the robot with a robotic voice. The model announces that it is going to take one marble out and asks whether the child is watching. Then the model demonstrates the inefficient strategy including two irrelevant actions (see Figure 2 A: moving the lever from one side to the other; B: pushing the button) and one relevant action (C: opening the flap covering the tube and inserting a magnetic rod to extract a marble). In Trial 2 (27 s), the model announces that it will get one more marble and performs the efficient strategy, only containing the relevant action C. In Trial 3 (47 s) the model repeats that it will get another marble and performs (as in Test trial 1) the inefficient strategy (A–C).[1]

Due to the robot's limited motor abilities, the critical actions looked somewhat atypical for humans. For example, the robot moved the lever by placing its palms horizontally under the lever, facing the ceiling and pushing the lever left and right with the side of the hand (Figure 2 A). To ensure comparability across conditions, the human model used motion patterns similar to those performed by the robot. Due to the robot's small size, it stood on the table during the action, whereas the human sat on a chair next to the table. To avoid mechanical background noise during filming, the soundtrack was recorded separately from the video and later added to critical segments of the video, like the self-introduction of the protagonist and the demonstration of the irrelevant actions (i.e., sound-matching the movement of the lever and the pushing of the button).

Procedure

Each child was invited to enter a quiet room with the experimenter. On average the procedure took approximately 17 min. First, there was a warm-up phase during which the child and the experimenter played a game of blowing cotton balls into a goal to win golden marbles. This phase served to familiarize the children with the experimenter and to explain that golden marbles could be exchanged for stickers. The first test trial began with the experimenter introducing the box ("Let's now look at this special box. In this box are more golden marbles which can be exchanged for stickers."). The child had a few seconds to visually inspect the box before it was covered again. The experimenter then said: "Before you have a go, please watch this video. While you are watching, I need to write down something on my sheet, but you watch carefully. Later, you will have a go." Then the experimenter turned away and sat down on a table turning her back to the child. Unlike in other oi studies (but see Hoehl et al., 2014 or Schleihauf et al., 2018) children saw only one demonstration per trial, to create a more natural learning context and to keep their attention high, as we presented a total of three trials.

After starting the video (on a laptop with a 33-cm display mounted in full view), the experimenter sat down at a near-by table to fill out a form. Following the video demonstration of Trial 1 (i.e., demonstration of the inefficient strategy), the experimenter explained that it would now be the child's turn to retrieve a marble. She explicitly stated that the child could accomplish this goal, however, he or she liked. In the next step, the experimenter uncovered the box and placed all material (i.e., box and magnetic rod) in front of the child. Then, she told the child that she needed to go outside for a moment. Hence, the child remained alone during the imitation phase. As soon as the child had successfully taken a marble out of the box or 2 min had elapsed, the experimenter returned and helped the child to choose a sticker from a sticker box in exchange for the marble. Then the box was covered again, and the next trial started. The experimenter explained that it was now the model's turn again before she started the presentation of the next video clip. Trials 2 and 3 were equivalent to Trial 1, but differed in terms of the video scenes presented. At the end of the experimental session the experimenter asked the child whether the model shown in the video was alive. Responses to the question could not be analyzed statistically, as several children did not answer, but will serve to interpret our data on a qualitative level in the discussion.

Children in the baseline condition experienced the same procedure with the following two differences: (a) they never saw a video demonstration, (b) they only participated in one single trial. As in the experimental conditions, the experimenter introduced the box and placed it in front of the child (together with the magnetic rod). Then the child was told that she could extract a marble from the box, however, she liked. Next, the experimenter left the room and waited either until the child extracted

[1]In comparison to the precursor studies using the same apparatus, we did not include non-contact inefficient actions like hand clapping or tapping the palm of the hand with the rod three times. These actions could not be programmed into the robot, and they were very rarely copied by children participating in earlier studies.

a marble or until 2 min had passed. This condition was included in the analysis to check if performance levels of irrelevant actions exceeded performance levels of these actions due to curiosity or exploration.

Coding and Reliability

We applied two different types of coding, an oi score and an eagerness score. The oi score informed us whether an irrelevant action was performed or not. The eagerness score entails information on the precision of imitative acts and the number of repetitions of each single action. Since the pattern of results obtained using data from each coding scheme did not differ, only report the results of the oi score here. The results of the eagerness score can be found in Supporting Information.

For the oi score, we first applied a binary coding (0 = action not performed, 1 = action performed) for each action. Pushing the lever was coded as performed as soon as the lever was pushed in one direction. Pushing the button was coded as performed as soon as the button was pushed down (no matter if this action was performed with the rod or with the finger). We then added up these values per test trial to calculate the oi score that indicated whether a child performed none, one (lever or button) or both (lever and button) irrelevant actions (e.g., lever pushed = 1 and button pushed = 1 resulted in an oi score of 2).

Scores were coded from video recordings of children's behavior during each imitation test trial. The primary coder was a research assistant who was blind to the hypotheses of the study. A second coder, also blind to the hypotheses, rated 25% of the videos based on edited versions showing only the child operating the box. The inter-rater reliability (Cohen's κ) was excellent (reliability of 1.0), suggesting reliable assessments of oi scores.

Statistical Analysis

To answer our research questions, we applied two main logistic Generalized Linear Mixed Models fitted via maximum likelihood (GLMMs; Baayen, 2008) using the statistical program R (version 3.4.3; R Core Team, 2015) together with the function "glmer" of the package "lme4" (Bates, Maechler, Bolker, & Walker, 2015). Our first research question was whether 5- to 6-year-olds over-imitate based on a social goal, and whether they consequently flexibly switch back to an inefficient strategy even after having performed a more efficient strategy? To answer this question, we need to compare oi rates in the experimental conditions with the spontaneous baseline performance of irrelevant actions. Thus, in GLMM 1, data of both experimental conditions (human and robot condition) as well as the baseline condition were analyzed. Second, we wanted to know, whether children in this age range were equally willing to learn an inefficient strategy from a humanoid robot as they are from a human experimenter. To answer this question, in GLMM 2, we analyzed only data from both experimental conditions.

We decided to analyze our data using GLMMs, because our data did not fulfil the assumptions for a classical analysis of variance (ANOVA). An ANOVA assumes normally distributed and homogeneous residuals. These assumptions cannot be met for data that has an upper bound (our data were based on a binary coding scheme and the summarizing oi score was bound between 0 and 2). In such a case, moving toward a mixed-modelling framework is suggested (Jaeger, 2008; Kristensen & Hansen, 2004). Details regarding this statistical method, assumption tests as well as the exact model equations for all the GLMMs can be found in Supporting Information.

GLMM 1

In GLMM 1, we investigated whether oi rates exceeded baseline level across different test trials. Thus, we included the data of both experimental conditions (human condition and robot condition) and the baseline condition into the analysis. We focused on differences in children's responses in all three trials (1, 2, 3) and on the differences in children's responses between conditions (i.e., human vs. robot). Since we were mainly interested in whether the two conditions differ in Trial 3, we included the interaction between test trial and condition into the model equation. It should be noted that we had only one trial for the baseline testing, however, the statistical model is treating our data as if we had three baseline trials with identical outcomes. This is an acceptable method because it allows us to compare each trial in both conditions against the same baseline trial. However, looking at the effects of other predictors (such as child's sex) would falsify the results, because then the multiplied data points from the baseline condition would count as actual measurements. Thus child's sex was not included as a predictor in this model. To account for repeated measurements, we also included random effects (individual's identity) and random slopes (test trials within individual's identity). A detailed description of this analysis can be found in Supporting Information. To account for multiple testing issues, we first tested the overall effect of all three test predictors and the included interaction. Therefore, we compared the full model's deviance with that of a null model comprising only random effects to examine whether the inclusion of the test predictors provided a better fit to the data than participant identity alone. The full model provided a significantly better fit to the data than the null model ($\chi^2 = 132.44$, $df = 8$, $p < .0001$). To determine the effects of each predictor we further compared the full model with the corresponding reduced models that lacked the predictor of interest.

Since we found a significant interaction between the predictors condition and test trial, we inspected the baseline differences in more detail. Therefore, we performed six additional post hoc GLMMs to compare oi scores for each trial as a function of experimental conditions (robot model, human model) with the baseline condition (no demonstration) to determine whether children actually showed oi in individual trials. Each post hoc GLMM included the Baseline Condition as well as one test trial of each condition. To avoid inflating the Type I error rate, a Holm–Bonferroni correction was applied for each post hoc comparison.

GLMM 2

In GLMM 2, we included only the data of the experimental conditions to have a direct and easy to interpret comparison between the human and robot condition. Thus, GLMM 2 informed us whether children's imitative responses differed following the demonstration of the human or the robot model. The baseline data were not included this time. Thus, we could add other variables of interest. The model equation for GLMM 2 included test trial and condition and their interaction as predictors. Furthermore, we now considered child's sex and the interaction between child's sex and condition as potential predictor of children's imitative performance for two reasons: (a) Previous studies revealed gender effects showing that boys over-imitate more than girls (Frick, Clément, & Gruber, 2017; Schleihauf et al., 2018); (b) it seemed possible that both gender groups respond differently to robots and humans. Again, we included random effects (individual's identity) and random slopes (test trials within individual's identity).

First, we tested the overall effect of all test predictors and the included interactions by comparing the full model's deviance with that of a null model comprising only random effects. The full model provided a significantly better fit to the data than the null model ($\chi^2 = 67.528$, $df = 7$, $p < .0001$). Then, to determine the effects of each predictor we further compared the full model with the corresponding reduced models that lacked the predictor of interest. However, neither the interaction between test trial and condition ($\chi^2 = 0.393$, $df = 2$, $p = .822$) nor the interaction between sex and condition ($\chi^2 = 1.392$, $df = 1$, $p = .238$) reached statistical significance. Thus, we dropped these terms from the model and fitted a second model only comprising the main effects of the predictors. We again checked, that this model explained the data better than a null model, comprising only the random effect and random slope ($\chi^2 = 65.82$, $df = 4$, $p < .0001$). To check whether single main effects reached statistical significance we again performed comparisons between the main-effects model and respective reduced models, with the reduced models lacking the predictor of interest. All statistical parameters for GLMM 2 reported in the tables are from the model only comprising the main effects.

Results

GLMM 1

GLMM 1, which included data from both experimental conditions (human & robot) and from the baseline condition, revealed a significant effect for the interaction of test trial and condition (see Table 2). The data and the model lines for the point estimates of GLMM 1 are depicted in Figure 3. For a detailed report of estimates, standard errors and confidence intervals, please have a look in Supporting Information (Table S4). The equivalent analysis with the eagerness score revealed the same pattern of results (Table S10).

Baseline Comparisons

Post hoc comparisons with the baseline condition (Table 2; Table S5) revealed that in the human condition and in the robot condition children performed significantly more irrelevant actions than children in the baseline condition in Test trials 1 and 3, whereas this was not the case for Trial 2. Baseline comparisons with the eagerness score revealed the same pattern of results (Table S11).

GLMM 2

For GLMM 2, which only included data from both experimental conditions (human & robot), the

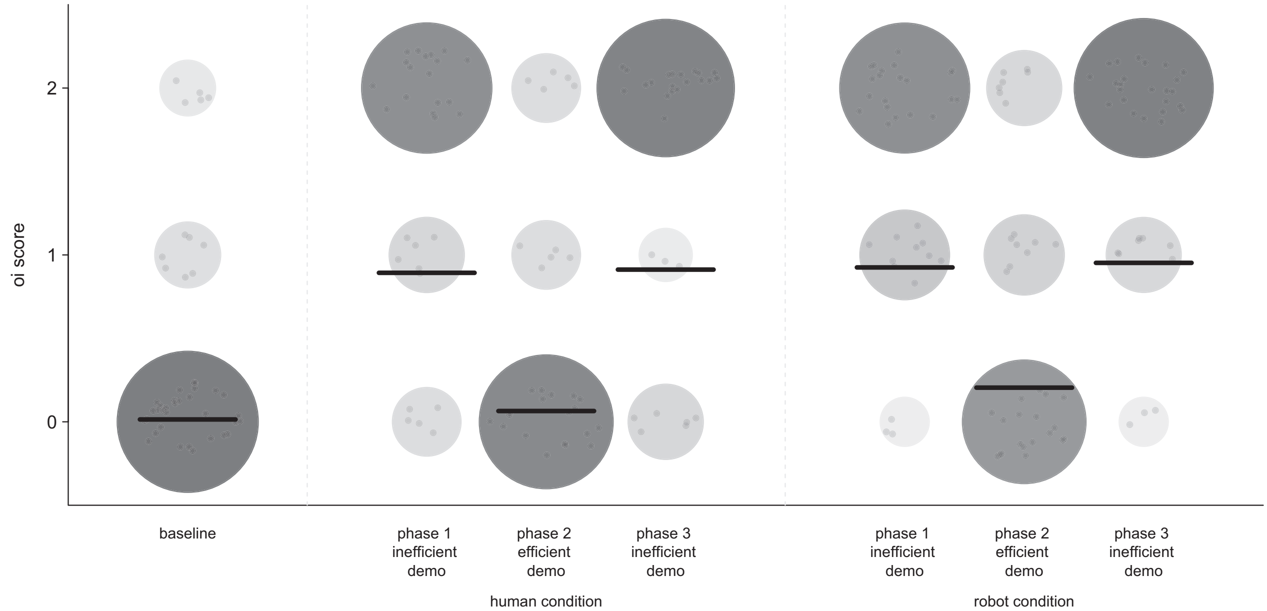

Figure 3. Oi scores as a function of the predictors test trial and condition. The frequency of the oi scores is represented by the size of the large circles as well as by the number of small circles (each small circle represents one data point). Lines represent the point estimates of the Generalized Linear Mixed Model 1.

comparisons between the full and reduced models revealed significant main effects for test trial and sex (see Table 3). Detailed reports of estimates, standard errors and confidence intervals can again be found in Supporting Information (Table S6). Descriptive analyses revealed the following changes in oi: From Trial 1 (demonstration of inefficient strategy) to Trial 2 (demonstration of only efficient strategy) the average oi rate decreased, and from Trial 2 to Trial 3 (demonstration of inefficient strategy) it increased again (see Figure 3 and Table 4).

Table 2
Results of the likelihood ratio tests for the comparisons between the full model of GLMM 1 and the reduced model lacking the predictors of interest, and the post-hoc baseline comparisons

GLMM 1	χ^2	*df*	*p*
Test trial: condition	32.205	4	< .0001*
Post hoc baseline comparisons with			
Human condition—Test trial 1	37.52	1	< .0001*
Human condition—Test trial 2	.119	1	.730
Human condition—Test trial 3	47.099	1	< .0001*
Robot condition—Test trial 1	42.253	1	< .0001*
Robot condition—Test trial 1	0.878	1	.349
Robot condition—Test trial 3	53.344	1	< .0001*

Note. For baseline comparisons critical α is Holm-Bonferroni-corrected, α′ = .0083. GLMM = Generalized Linear Mixed Model. For GLMM 1: *$p < .05$, for baseline comparisons: *$p < .0083$ (*p*-value adjusted for multiple testing).

The effect of condition (human vs. robot) was not significant, but the effect sex reached level of significance (see Table 3). Boys more often produced irrelevant actions than girls did ($M_{boys} = 1.35$, $M_{girls} = 1.07$), which is consistent with previous findings (Frick et al., 2017). No comparable gender difference was found for the baseline condition ($\chi^2 = 0.157$, $df = 4$, $p = .692$), which would have been an indication that boys find the apparatus simply more interesting than girls.

When these analyses were run with the eagerness score as dependent variable, results revealed a highly similar pattern of results, with only one exception: the observed sex differences could no longer be found. (see Table S12 for more details).

Table 3
Results of the likelihood ratio tests for the comparisons between the full model of GLMM 2 and the reduced models lacking the predictors of interest

GLMM 2	χ^2	*df*	*p*
Condition	1.203	1	.273
Test trial	60.029	2	< .0001*
Sex	4.232	1	.040*

Note. GLMM = Generalized Linear Mixed Model. For GLMM 2: *$p < .05$.

Table 4

Descriptive Information: Number of Children Who Re-Enacted Each of the Two Irrelevant Actions, Mean oi Score and Standard Deviation for Each Test Trial of Each Condition

	Frequency of each nonfunctional action performed		Frequencies of oi scores			
Condition	Pushing lever	Pushing button	0	1	2	Mean oi score (*SD*)
Baseline (*N* = 44)	6	11	32	7	5	0.39 (.69)
Human condition (*n* = 29)						
Test trial 1	24	18	5	6	18	1.45 (.78)
Test trial 2	9	6	19	5	5	0.52 (.78)
Test trial 3	22	21	6	3	20	1.48 (.83)
Robot condition (n = 34)						
Test trial 1	31	21	3	10	21	1.53 (.66)
Test trial 2	12	10	19	8	7	0.65 (.81)
Test trial 3	27	28	3	7	24	1.62 (.95)

Discussion

The aim of this study was to shed light on the basic functions and mechanisms of the human tendency to copy irrelevant actions, often labeled as oi. We compared children's production of irrelevant actions (a) before and after they were informed about the efficient strategy to solve the problem at hand and (b) when the target actions were demonstrated by a human or a robot model. Five-to-6-year-olds were randomly assigned to watch videos of either a human or a robot model retrieving a marble from a puzzle box. Both models moved in a similar way and expressed communicative and ostensive cues, thus supporting the impression of a social-interactive partner. Children saw three demonstrations on three separate trials: (1) an inefficient strategy, (2) an efficient strategy, and (3) the inefficient strategy again. Following each demonstration, the children were left alone and allowed to solve the task on their own. Results revealed that preschoolers matched their imitative response across all three trials with the respective demonstration by the human or the robot model, even though they only watched the actions on video and were left alone during the test trial. Comparisons with a baseline condition (i.e., no demonstration) confirmed that children over-imitated following the first demonstration of the inefficient strategy (Trial 1), then switched to an efficient strategy following the efficient demonstration (Trial 2) and switched back to an inefficient solution after the model had demonstrated the inefficient strategy once more (Trial 3). Thus, children adopted the inefficient strategy in Test trials 1 and 3, but used the efficient strategy in Test trial 2. We did not find any differences in children's oi when they observed a human versus a robot model. Furthermore, we found that boys over-imitated more often than girls (only when oi score, not the eagerness score, served as dependent measure).

Imitation Performance Across Trials

Our first research question was whether young children can switch back and forth between the inefficient and the efficient strategy or whether they primarily follow an instrumental goal, using the efficient strategy as soon as they learn a direct way to attain a reward in Trial 2. The findings reported for Trial 1 are in line with a large number of oi studies reporting that children of the tested age range copy an inefficient strategy from a communicative, ostensive model (e.g., Hoehl et al., 2014; Lyons et al., 2011; McGuigan & Whiten, 2009). However, this finding does not reveal whether children over-imitated for social reasons (i.e., based on a normative orientation or a motivation to affiliate with the experimenter) or because they assumed that all actions were indeed necessary to retrieve the reward (i.e., based on false causal reasoning). Findings for Trial 2 are also in line with previous results showing that a demonstration of the efficient strategy reduces oi rates (Hoehl et al., 2014; Schleihauf et al., 2018). Differing from previous work, we introduced a third test trial. Here, we expected participants to experience a conflict between the instrumental motivation to simply obtain the reward in an efficient way and the social motivation to affiliate with the model or to follow the social norm "I will do what you do." In previous studies using a similar puzzle box and similar tools, children who first saw and applied the efficient strategy and then watched the demonstration of the inefficient strategy, did not show oi rates exceeding baseline level (Schleihauf et al., 2018). Hence, they seem well able to understand that the irrelevant actions are not really necessary to get the reward and should have learned this lesson during Trial 2. In the present case, however, children switched back to an inefficient strategy in Test trial 3, despite having previously watched and performed a more efficient strategy in Test trial 2. The finding speaks against the idea that oi is exclusively based on instrumental learning. Rather, children were found to be highly flexible in their imitative responses (see also Williamson & Meltzoff, 2011; Williamson et al., 2008).

But does this imply that their underlying motivation was social in nature? In principle, children may have thought that the inefficient actions were actually needed to get the marble out on Trials 1 and 3 but not on Trial 2. This only makes sense if they assumed that the causal logic of the task changed across trials. Although the puzzle box was covered with a blanket while the children watched the videos, it was always in full view and the children saw that it was neither manipulated nor replaced during the experimental session. For this reason, we can hardly attribute changes in children's behavior across trials to (assumed) changes regarding the box or the mechanism leading to the retrieval of the reward.

Hence, it seems highly unlikely that automatic causal encoding (Lyons et al., 2007, 2011) provided the main reason for the high oi rates in Trial 3. Rather, children may have wanted to either affiliate with the model demonstrating the actions (Nielsen & Blank, 2011) or to follow social norms (Kenward, 2012; Kenward et al., 2011; Keupp, Behne, Zachow, Kasbohm, & Rakoczy, 2015), thus acting based on a social motivation. In line with this interpretation, the children's responses could result from an affiliation with the model or from applying a rather simple rule: "I will do what you do." Anecdotal observations support this idea: Some children participating in Trial 3 first took the reward out of the box using the efficient strategy (thus revealing functional knowledge), but then put it back inside the box again in order to perform the inefficient strategy. Some also reached out to take the marble from the box, then stopped themselves in the middle of this (efficient) action only to perform both irrelevant actions before repeating the efficient action (see also Williamson et al., 2008). Such observations indicate that children made a conscious choice between using the efficient or the inefficient strategy, and then decided against using the efficient strategy.

Importantly, the model in our study stayed the same throughout all three trials whereas it typically changed in previous studies using a similar paradigm (e.g., Hoehl et al., 2014; Nielsen & Blank, 2011). For that reason we speculate that children who saw the model first perform the inefficient strategy on Trial 1, and then the efficient strategy during Trial 2 may have reasoned that the model was well aware of the most efficient strategy to get the marble even during Trial 1 (see also Nielsen et al., 2015). In Trial 3, the model again used the inefficient strategy. In an attempt to explain this behavior, children may have assumed that the model "has a good reason to do it that way," thus becoming engaged in "rational imitation" (Gergely, Bekkering, & Király, 2002). Whether these "good reasons" could be instrumental or a social remains unclear. However, whether or not this is an example of "rational imitation" and what mental attributions the children make about the robot need to be investigated in future research.

Imitating a Human Versus a Robot Model

With regard to our second research question, we tested whether or not oi is restricted to interactions within the human species. When comparing oi performances between the respective human and robot conditions, we found a striking similarity between children's responses in both conditions across all three test trials (see Figure 3). Interestingly, the robotic voice did not influence children's willingness to imitate in our study, although previous studies found that children's imitation rates were reduced when the demonstration was performed by a model with a foreign accent (Howard, Henderson, Carrazza, & Woodward, 2015).

The finding that children not only over-imitated the human model but also the robot model in Test trial 3 (i.e., when they already knew that there was a more efficient solution) was most remarkable. As argued before, performance in test trial 3 is likely to result from processes of social reasoning and a social motivation, thus raising the interesting question why we did not find any difference in oi between the human and the robot condition. At first sight this seems to indicate that children were equally willing to accept a human or a robot as their role model.

We used a robot with a human body configuration, including hands to manipulate tools, facial features including eyes, speaking to the children in their mother tongue. All these different attributes are known to support the identification of social agents and animate beings (e.g., Pauen, 2000) and influence children's readiness to learn cultural practices (Gergely & Csibra, 2006). In this study, we observed social responses to the videos in some children who waved back to the robot in the introductory greeting scene which suggest that they actually perceived the robot as a social-interactive partner and might not understand that robots do not have a mind (Kahn et al., 2012; Sommer et al., 2019).

In line with this idea, earlier findings demonstrate children's willingness to accept humanoid robots as interactive guides (Meltzoff et al., 2010; Shiomi et al., 2006), peer tutors (Kanda et al., 2004; Meltzoff et al., 2009; Tanaka et al., 2007) or even as therapist (Warren et al., 2015). Further evidence supporting this view comes from work showing that children faithfully copied actions from humanoid robots (Itakura et al., 2008) or puppets (McGuigan & Robertson, 2015) as models in imitation tasks.

In this study, when children were asked whether they thought the model was alive or not, they revealed some uncertainty in their judgment: In the human condition, 62.5% responded "yes," 15.6% responded "no," 6% stated "I don't know" and the rest (15.9%) did not answer at all. In the robot condition, 33% answered with "yes," 54.5% answered with "no" and 12.5% stated that they "do not know." This could either be due to a lack of conceptual understanding of the distinction between living and nonliving entities, to general problems interpreting the living-status of entities presented on video, to language problems in understanding the question, or to the overlap in static and dynamic features between the human and the robot model presented. Interestingly, also Sommer et al. (2019) found that children tend to ascribe mental life to humanoid robots. However, more work is still needed that systematically varies characteristics of both kinds of protagonists in order to shed more light on the circumstances that lead young children to treat robots as social-interactive partners.

To explore in more detail what kind of behavior and which characteristics of the model might evoke oi, future studies should systematically manipulate the static and dynamic features of the robot. For example, it would be interesting to leave out the first greeting scene, as this may help children establish a social relationship with the protagonist. In addition, it would be worthwhile to present a robot without facial features, as eyes are known to be critical for identifying social interactive partners. Both manipulations should reduce imitation rates in the robot condition.

Although we favor the explanation that human children attribute social characteristics to the robot, alternative interpretations also need to be considered. They refer to (a) the social role of the experimenter, (b) the nature of the demonstrated action, (c) untypical movements of the human model, and (d) the video demonstration. These factors address possible limitations of the study and should be carefully considered in future studies:

The Social Role of the Experimenter

In this study, a live-experimenter first played a warm-up game with the children, then introduced the laptop and guided the child through the experimental session. One could thus argue that children perceived the given task as a social game with the experimenter. Alternatively, children may also have perceived the experimenter as a high-status person (McGuigan, 2013) who expected them to follow instructions (Kenward, 2012). In either case, imitation fidelity in both experimental conditions could be attributed to a social motivation of the children to engage with the live-experimenter rather than with the model on the screen, thus marginalizing the impact of the models and the differences between conditions.

Though we cannot rule out this interpretation, we would like to point out that the experimenter turned away from the video screen while the children were watching the demonstration (i.e., she was pretending to write down something sitting at a desk and turning her back on the scene), and that she left the room during the testing phase. This implies that the children were left alone during the test phase and free to retrieve a marble any way they liked. Because it seemed impossible to run an experimental task involving young children without having an experimenter to interact with the child, we thus tried to minimize social pressure on the child and prevented children from socially interacting with the experimenter during the critical demonstration and imitation phase of the experimental session.

Furthermore, it seems noteworthy that previous work on imitation fidelity also used video demonstrations provided by a live experimenter, but nonetheless found systematic differences between experimental conditions (e.g., DiYanni, Corriveau, Kurkul, Nasrini, & Nini, 2015; Watson-Jones, Legare, Whitehouse, & Clegg, 2014). We therefore conclude that the lack of any significant difference in oi rates between the human and the robot condition in our study is unlikely to result from the presences of a live-experimenter. At the same time, we concede that the social relationship established between the child and the experimenter may have had some effect on the child's motivation in the present task. This would support the view that children's oi in the present task is primarily based on social motivations. Further studies may address this question more directly, for example, by manipulating the social behavior of the experimenter and testing the effects of this manipulation on children's oi performance following video demonstrations.

The Nature of Demonstrated Actions

The demonstrated action sequence concluded in a goal-directed action (i.e., extracting the marble). Thus, it is possible that participants perceived the entire action sequence (including the inefficient actions) as functional rather than social, which led to similar rates of oi following the robot and the human demonstration. It would be interesting to see whether or not oi rates in both conditions remained similar in a study exploring children's tendency to copy nongoal-oriented versus goal-directed actions. This might be a fruitful avenue for future research.

Furthermore, we presented only one demonstration of the target actions on each trial to keep the entire procedure short. However, more demonstrations may enhance the social context and lead to differences between the robot and the human condition.

Untypical Movements of Human Model

In the human condition, we used a female model and matched her movements with those of the robot model in order to keep both videos comparable. As a result, some movements of the person looked mechanical and rather unnatural. Such mechanical movements may have confused some children, thus reducing their oi rates in the person condition and this may have resulted in an underestimation of group differences in oi rates between the robot and the human conditions. Thus, it is suggested for future research to add an additional condition with natural movements.

Video Demonstrations

In this study, children only saw video instead of live demonstrations of both models. It is possible that, the differences between the human and the robot model were reduced because the presentation medium was the same in both conditions and might have been more pronounced in a live situation. Thus, it should be envisaged to implement live demonstrations in future studies.

Sex Differences

Some earlier work also found that boys show more oi than girls (Frick et al., 2017; Schleihauf, Pauen, & Hoehl, 2019). Since boys did not explore the apparatus and the tools more than girls did in the baseline condition (but see Schleihauf et al., 2019), it seems unlikely that in this study the puzzle box task was generally more attractive for boys. Frick et al. (2017) relate the boys' stronger preference to imitate irrelevant actions to the finding that boys use more tools than girls when solving a problem (Gredlein & Bjorklund, 2005). This may lead to higher attention to objects (such as the lever or the button in our study) in boys than in girls, which—in consequence—could have increased oi rates in the present task context. Importantly, child's sex did not interact with condition, so there were no gender differences in children's preference to copy the human or the robot model.

Summary and Implications

Taken together, this study demonstrates that children are willing to perform causally irrelevant actions even when they know a more efficient strategy to solve the problem at hand. Our data thus demonstrate that an instrumental motivation may not always dominate responses in social settings, even with limited social pressure and a clear instrumental goal at hand. Rather, social motivations seem to play an important role in this context as well.

We further conclude that the high oi rates observed across trials, the social signals given by children while watching the video, and the lack of any difference in children's responses between the robot and the human condition jointly lead to the impression that the Nao-robot has been perceived as a social agent.

These findings give rise to the hypothesis that oi may not provide a learning mechanism restricted to the transmission of culturally relevant knowledge exclusively within the human species. Artificial intelligence is an integral part of our present culture. Before long, more children will encounter robots as teachers in educational settings. To prepare for the future, we should increase our knowledge about similarities and differences in observational learning from and imitation of robot models. Future studies addressing this issue are clearly needed.

Ethical Statement

We received ethical clearance from the faculty's local ethics committee. Furthermore, we received full written informed consent from the children's parents.

Data Availability Statement

Data and videos associated with this manuscript can be accessed at https://doi.org/10.17632/r96vch87fj.3.

References

André, V., Jost, C., Hausberger, M., Le Pévédic, B., Jubin, R., Duhaut, D., & Lemasson, A. (2014). Ethorobotics applied to human behaviour: Can animated objects influence children's behaviour in cognitive tasks? *Animal Behaviour*, *96*, 69–77. https://doi.org/10.1016/j.anbehav.2014.07.020

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge, UK: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Breazeal, C., Harris, P. L., DeSteno, D., Westlund, J. M. K., Dickens, L., & Jeong, S. (2016). Young children treat robots as informants. *Topics in Cognitive Science*, *8*, 481–491. https://doi.org/10.1111/tops.12192

Clay, Z., & Tennie, C. (2017). Is overimitation a uniquely human phenomenon? Insights from human children as compared to Bonobos. *Child Development*, *89*, 1535–1544. https://doi.org/10.1111/cdev.12857

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1149–1157. https://doi.org/10.1016/j.actpsy.2006.09.007

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology*, *47*, 1499–1503. https://doi.org/10.1037/a0025659

DiYanni, C. J., Corriveau, K. H., Kurkul, K., Nasrini, J., & Nini, D. (2015). The role of consensus and culture in children's imitation of inefficient actions. *Journal of Experimental Child Psychology*, *137*, 99–110. https://doi.org/10.1016/j.jecp.2015.04.004

Flynn, E., & Smith, K. (2012). Investigating the mechanisms of cultural acquisition—How pervasive is overimitation in adults? *Social Psychology*, *43*, 185–195. https://doi.org/10.1027/1864-9335/a000119

Frick, A., Clément, F., & Gruber, T. (2017). Evidence for a sex effect during overimitation: boys copy irrelevant modelled actions more than girls across cultures. *Royal Society Open Science*, *4*, 170367. https://doi.org/10.1098/rsos.170367

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*, 755. https://doi.org/10.1038/415755a

Gergely, G., & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In N. J. Enfield & S. C. Levenson (Eds.), *Roots of human sociality: Culture, cognition, and human interaction* (pp. 229–255). Oxford, UK: Berg.

Gredlein, J. M., & Bjorklund, D. F. (2005). Sex differences in young children's use of tools in a problem-solving task. *Human Nature*, *16*, 211–232. https://doi.org/10.1007/s12110-005-1004-5

Hoehl, S., Keupp, S., Schleihauf, H., McGuigan, N., Buttelmann, D., & Whiten, A. (2019). "Over-imitation": A review and appraisal of a decade of research. *Developmental Review*, *51*, 90–108. https://doi.org/10.1016/j.dr.2018.12.002

Hoehl, S., Zettersten, M., Schleihauf, H., Grätz, S., & Pauen, S. (2014). The role of social interaction and pedagogical cues for eliciting and reducing overimitation in preschoolers. *Journal of Experimental Child Psychology*, *122*, 122–133. https://doi.org/10.1016/j.jecp.2013.12.012

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, *8*, 164–181. https://doi.org/10.1007/s10071-004-0239-6

Howard, L. H., Henderson, A. M. E., Carrazza, C., & Woodward, A. L. (2015). Infants' and young children's imitation of linguistic in-group and out-group informants. *Child Development*, *86*, 259–275. https://doi.org/10.1111/cdev.12299

Itakura, S., Ishida, H., Kanda, T., Shimada, Y., Ishiguro, H., & Lee, K. (2008). How to build an intentional android: Infants imitation of a robot's goal-directed actions. *Infancy*, *13*, 519–532. https://doi.org/10.1080/15250000802329503

Jaeger, F. T. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jones, S. S. (2009). Imitation and empathy in infancy *Cognition, Brain, Behavior*, *8*, 391–413. Retrieved from https://www.researchgate.net/publication/266588462

Kahn, Jr., P. H., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., Shen, S., . . .Severson, R. L. (2012). *Do people hold a humanoid robot morally accountable for the harm it causes?* The seventh annual ACM/IEEE international conference (pp. 33–40). New York, NY: ACM. https://doi.org/10.1145/2157689.2157696

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human Computer Interaction*, *19*, 61–84. https://doi.org/10.1207/s15327051hci1901&2_4

Kenward, B. (2012). Over-imitating preschoolers believe unnecessary actions are normative and enforce their performance by a third party. *Journal of Experimental Child Psychology*, *112*, 195–207. https://doi.org/10.1016/j.jecp.2012.02.006

Kenward, B., Karlsson, M., & Persson, J. (2011). Over-imitation is better explained by norm learning than by distorted causal learning. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 1239–1246. https://doi.org/10.1016/j.neuroimage.2005.05.033

Keupp, S., Behne, T., & Rakoczy, H. (2013). Why do children overimitate? Normativity is crucial. *Journal of Experimental Child Psychology*, *116*, 392–406. https://doi.org/10.1016/j.jecp.2013.07.002

Keupp, S., Behne, T., Zachow, J., Kasbohm, A., & Rakoczy, H. (2015). Over-imitation is not automatic: Context sensitivity in children's overimitation and action interpretation of causally irrelevant actions. *Journal of Experimental Child Psychology*, *130*, 163–175. https://doi.org/10.1016/j.jecp.2014.10.005

Kristensen, M., & Hansen, T. (2004). Statistical analyses of repeated measures in physiological research: a tutorial. *Advances in Physiology Education*, *28*, 2–14. https://doi.org/10.1152/advan.00042.2003

Legare, C. H. (2017). Cumulative cultural learning: Development and diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 7877–7883. https://doi.org/10.1073/pnas.1620743114

Legerstee, M. (1992). A review of the animate-inanimate distinction in infancy: Implications for models of social and cognitive knowing. *Early Development and Parenting*, *1*, 59–67. https://doi.org/10.1002/edp.2430010202

Lyons, D. E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil, F. C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1158–1167. https://doi.org/10.1098/rstb.2010.0335

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 19751–19756. https://doi.org/10.1073/pnas.0704452104

McGuigan, N. (2013). The influence of model status on the tendency of young children to over-imitate. *Journal of Experimental Child Psychology*, *1–8*, https://doi.org/10.1016/j.jecp.2013.05.004

McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying: Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology*, *102*, 1–18. https://doi.org/10.1348/000712610X493115

McGuigan, N., & Robertson, S. (2015). The influence of peers on the tendency of 3- and 4-year-old children to over-imitate. *Journal of Experimental Child Psychology*, *136*, 42–54. https://doi.org/10.1016/j.jecp.2015.03.004

McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. *Cognitive Development*, *22*, 353–364. https://doi.org/10.1016/j.cogdev.2007.01.001

McGuigan, N., & Whiten, A. (2009). Emulation and "overemulation" in the social learning of causally opaque versus causally transparent tool use by 23- and 30-month-olds. *Journal of Experimental Child Psychology*, *104*, 367–381. https://doi.org/10.1016/j.jecp.2009.07.001

Meltzoff, A. N., Brooks, R., Shon, A. P., & Rao, R. P. N. (2010). "Social" robots are psychological agents for infants: A test of gaze following. *Neural Networks*, *23*, 966–972. https://doi.org/10.1016/j.neunet.2010.09.005

Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science*, *325*, 284–288. https://doi.org/10.1126/science.1175626

Meltzoff, A. N., & Marshall, P. J. (2018). Human infant imitation as a social survival circuit. *Current Opinion in Behavioral Sciences*, *24*, 130–136. https://doi.org/10.1016/j.cobeha.2018.09.006

Nielsen, M., & Blank, C. (2011). Imitation in young children: When who gets copied is more important than what gets copied. *Developmental Psychology*, *47*, 1050–1053. https://doi.org/10.1037/a0023866

Nielsen, M., Kapitány, R., & Elkins, R. (2015). The perpetuation of ritualistic actions as revealed by young children's transmission of normative behavior. *Evolution and Human Behavior*, *36*, 191–198. https://doi.org/10.1016/j.evolhumbehav.2014.11.002

Nielsen, M., & Tomaselli, K. (2010). Overimitation in Kalahari bushman children and the origins of human cultural cognition. *Psychological Science*, *21*, 729–736. https://doi.org/10.1177/0956797610368808

Nielsen, M., Tomaselli, K., & Kapitány, R. (2018). The influence of goal demotion on children's reproduction of ritual behavior. *Evolution and Human Behavior*, *39*, 343–348. https://doi.org/10.1016/j.evolhumbehav.2018.02.006

O'Connell, L., Poulin-Dubois, D., Demke, T., & Guay, A. (2009). Can infants use a nonhuman agent's gaze direction to establish word-object relations? *Infancy*, *14*, 414–438. https://doi.org/10.1080/15250000902994073

Over, H., & Carpenter, M. (2012). Putting the social into social learning: explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*, *126*, 182–192. https://doi.org/10.1037/a0024555

Pauen, S. (2000). Early differentiation within the animate domain: Are humans something special? *Journal of Experimental Child Psychology*, *75*, 134–151. https://doi.org/10.1006/jecp.1999.2530

R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rakison, D. H., & Poulin-Dubois, D. (2001). Developmental origin of the animate-inanimate distinction. *Psychological Bulletin*, *127*, 209–228. https://doi.org/10.1037//0033-2909.127.2.209

Schleihauf, H., Graetz, S., Pauen, S., & Hoehl, S. (2018). Contrasting social and cognitive accounts on overimitation: The role of causal transparency and prior experiences. *Child Development*, *89*, 1039–1055. https://doi.org/10.1111/cdev.12780

Schleihauf, H., Pauen, S., & Hoehl, S. (2019). Minimal group formation influences on over-imitation. *Cognitive*

Development, *50*, 222–236. https://doi.org/10.1016/j.cogdev.2019.04.004

Shiomi, M., Kanda, T., Ishiguro, H., & Hagita, N. (2006). Interactive humanoid robots for a science museum. *Proceeding of the 1st ACM SIGCHI/SIGART conference* (pp. 305–312). New York, NY: ACM. https://doi.org/10.1145/1121241.1121293

Sommer, K., Nielsen, M., Draheim, M., Redshaw, J., Vanman, E. J., & Wilks, M. (2019). Children's perceptions of the moral worth of live agents, robots and inanimate objects. *Journal of Experimental Child Psychology*, *187*, 104656. https://doi.org/10.1016/j.jecp.2019.06.009

Tanaka, F., Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 17954–17958. https://doi.org/10.1073/pnas.0707769104

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Science*, *28*, 1–62. https://doi.org/10.1017/S0140525X05000129

Warren, Z., Zheng, Z., Das, S., Young, E. M., Swanson, A., Weitlauf, A., & Sarkar, N. (2015). Brief Report: Development of a robotic intervention platform for young children with ASD. *Journal of Autism and Developmental Disorders*, *45*, 3870–3876. https://doi.org/10.1007/s10803-014-2334-0

Watson-Jones, R. E., Legare, C. H., Whitehouse, H., & Clegg, J. M. (2014). Task-specific effects of ostracism on imitative fidelity in early childhood. *Evolution and Human Behavior*, *35*, 204–210. https://doi.org/10.1016/j.evolhumbehav.2014.01.004

Whiten, A., Allan, G., Devlin, S., Kseib, N., Raw, N., & McGuigan, N. (2016). Social learning in the real-world: "over-imitation" occurs in both children and adults unaware of participation in an experiment and independently of social interaction. *PLoS One*, *11*, e0159920. https://doi.org/10.1371/journal.pone.0159920

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 2417–2428. https://doi.org/10.1111/j.1467-9280.1996.tb00386.x

Williamson, R. A., & Meltzoff, A. N. (2011). Own and others' prior experiences influence children's imitation of causal acts. *Cognitive Development*, *26*, 260–268. https://doi.org/10.1016/j.cogdev.2011.04.002

Williamson, R. A., Meltzoff, A. N., & Markman, E. M. (2008). Prior experiences and perceived efficacy influence 3-year-olds' imitation. *Developmental Psychology*, *44*, 275–285. https://doi.org/10.1037/0012-1649.44.1.275

Wood, L. A., Harrison, R. A., Lucas, A. J., McGuigan, N., Burdett, E. R. R., & Whiten, A. (2016). "Model age-based" and "copy when uncertain" biases in children's social learning of a novel task. *Journal of Experimental Child Psychology*, *150*, 1–13. https://doi.org/10.1016/j.jecp.2016.06.005

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2012). Context-dependent model-based biases in cultural transmission: children's imitation is affected by model age over model knowledge state. *Evolution and Human Behavior*, *33*, 387–394. https://doi.org/10.1016/j.evolhumbehav.2011.11.010

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Appendix S1. Detailed Description of the Statistical Method and Results