Age-Related Differences in Children's Spontaneous Gesturing with a Robot versus Human Instructor

Jason R. Wilson[1], Allison Langer[2], Lauren Howard[1], and Peter J. Marshall[2]

***Abstract*— Research on gestures in human-robot interaction has largely focused on finding that children may learn better and enjoy interacting with robots that gesture more often. However, no research to date has examined how children themselves spontaneously gesture in the presence of a human vs. robot instructor. A child's use of gesture might be indicative of engagement or rapport with the robot instructor and may provide key information about a robot instructor's efficacy or opportunities for intervention. As such, the current study examines 5-8-year-old children's rate of deictic and conventional gestures when being assisted by a robot vs. human instructor. Overall, we find age-related effects in children's gestures relation to the specific instructors. There is a significant negative correlation between age and gesture rate when learning from the human instructor, but no significant correlation with the robot instructor. These results are discussed in relation to children's perceptions of the instructor, task difficulty, and age-related cognitive development shifts.**

I. Introduction

Research has shown that gestures provide vital information for communication, above and beyond that conveyed by language (see [1]). Even early in development, children are sensitive to the information that gesture provides [2], and they utilize gestures to effectively communicate with others [3]. In the human-robot interaction (HRI) sphere, the study gesture has mainly explored how changes in the types of gestures robots produce or the rate of robot gestures can influence perceptions in humans (e.g., [4]). Although there is research exploring the influence of robotic gestures on children [], almost no work has been done exploring how cildren gesture spontaneously to robots. In the current study, we explore how 5- to 8-year old children spontaneously gesture to either a robot instructor and a human instructor during a learning paradigm.

Previous research is inconsistent concerning expected rates of gesturing across this age range. For example, studies that have employed paradigms meant to elicit increased gestures in human-human dyads have found that 10-year-olds gesture more than 6-year-olds [5]. However, research exploring a more spontaneous rate of gesture suggests a decrease in gesture as children approach age 7, potentially due to an increase in language sophistication that decreases the need for gesture [6].

Here, we are interested in understanding how children gesture toward a robot during a learning task. This work is novel in its focus on child-produced, rather than robot-produced, aspects of gesture during a human-robot interaction. We explore developmental changes in how often children spontaneously gesture with a robot vs. human instructor and what kinds of gesture they use. We consider 3 types of gestures: deictic (used to draw attention to an object in the environment), conventional (which represent concepts in a culture-specific manner), and representational (used to visually depict an attribute or actions). We find that children spontaneously use a variety of gestures when interacting with a robot, and that age accounts for some of the variation in the type and frequency of gestures children employ. Results from this study highlight the importance of integrating user gestures into the design of interactions involving robot instructors in learning contexts with young users. Furthermore, our coding scheme and results may inform the design of future studies investigating children's gestures with robots.

[1]Franklin & Marshall College, Lancaster, PA 17603, USA jrw@fandm.edu
[2]Temple University, Philadelphia, PA 19122, USA

II. Background

A. Child Gestures when Learning

Current gesture research in developmental psychology suggests that user gestures differ substantially between individuals [7]. While experimental studies remain scarce, learners' gesture rate may make a difference in learning, particularly in how learner gesture rate may be related to response to an instructor's gesture rate [7]–[9]. Given the utility of robots in educational contexts [10], [11], it is important to understand (a) gesture patterns by children toward a robot relative to a human, and (b) the implications of these gesture patterns for designing effective robot instructors.

B. Robot Gestures to Children in HRI

Studies have examined the effects of robots' ability to gesture based on user experience [12] For example, robot gesture has been shown to increase children's enjoyment and engagement in learning paradigms [13]–[15], and can also increase early imitation and learning [13]. Gesturing robots have effectively supported gesture learning for children with autism spectrum disorder who may demonstrate delays in gesture production and understanding [16]. Though gestures that emanate from robots have shown to be productive for early learning, a research focus on the child's own gestures in these contexts has been lacking.

C. User Gestures in HRI

Previous gesture work in HRI research has largely focused on equipping robots with advanced gesture recognition technology [17]–[19], For example, deictic gestures (e.g., pointing) have been used to disambiguate referring expressions in language [20], [21] and to improve the robot's understanding of user intent [22].

A few studies have investigated how user gestures may inform interaction effectiveness and quality. One study examined how adult participants used gestures to teach a robot a task and found that users rarely used deictic and symbolic gestures, relying mostly on object manipulations [23]. Another study attempted to prepare for automatic recognition of gestures by capturing a small dataset of how people naturally gesture [24]. Since no study has investigated child gestures towards a robot, and these gestures may be meaningful for successful interactions in learning contexts, the current study examines the association between age, child gesture rates, and learning in a tangram task with a human or robot instructor.

D. Age-related Trends in Child-Robot Interaction

Prior research has shown developmental trends in the ways children understand and interact with robots. For example, younger children trust a human more than a robot after playing with both partners [25], and younger children may trust inaccurate robots less than inaccurate humans, relative to older children [26]. Age may also contribute to how well they can attend and engage with a robot. Three-year-olds look less at a robot during an interaction than 4-year-olds [27], younger children, compared to older children, have been found to enjoy interacting with robots more, potentially because they are more prone to anthropomorphize robots [28]. In line with this work, we seek to uncover age-related trends in children's gesture rates towards a robot and human instructor.

III. Methods

A. Participants

The study involved a 1-hour visit to the investigators' research institute. Thirty-two children aged 5-8 years (M=6.81 years, SD=0. 94; 17 girls & 15 boys) completed the study. According to parent report, 97% of participants identified as White and 3% identified as "Other race/ethnicit". Seven children began the laboratory visit but were excluded from analyses due to failing to complete the puzzle task (n=6), or due to issues with the recording equipment (n=1). The Institutional Review Board (IRB) at the primary investigator's research institute reviewed and approved the study protocol. Participants' parents/legal guardians signed an online informed consent form, and children gave verbal assent before participating in the laboratory task.

Children were asked to complete two tangram puzzles, adapted from the TOSA Test of Spatial Ability [29]. Tangrams are a set of simply shaped pieces that can be arranged to resemble an object (e.g., cat, rabbit, boat). Their use has been found to tap spatial and mathematical abilities in children [30].

B. Tangram Task

Participants entered a testing room and were seated at a table across from either a human or robot instructor

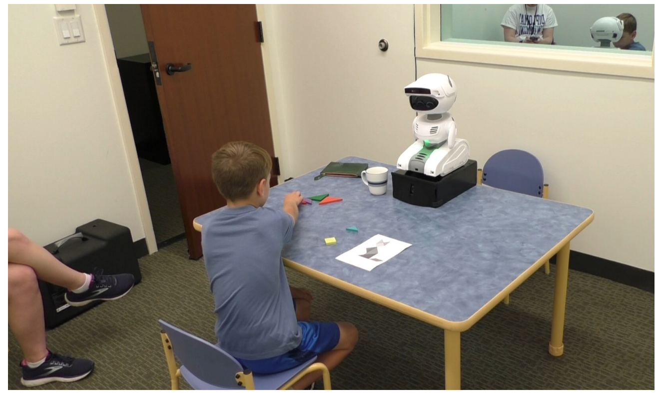

Fig. 1. A participant sits across from the robot instructor while working on the tangram task. A research assistant idly sits in the corner. Behind the one-way mirror is the robot controller.

(see Figure 1). Children interacted with both instructors, with the order of the first instructor counterbalanced across participants. A one-way mirror was behind the instructor. Three cameras recorded each interaction. Figure 1 shows the first camera angle, the black box under the robot in the figure contains the second camera, and a third was mounted on the wall above the mirror.

Tangram pieces and an illustration showing a target puzzle arrangement were on the table in front of the participant, along with a mug and a small green bag that were used to hide a piece of the puzzle. The research assistant sat in a chair behind the child and appeared busy so as not to influence the participant-instructor interactions.

For each instructor type, participants first completed an introduction phase to orient the child to the instructor and the task. The instructor asked the child warm up questions and taught them how to complete a 2-shape tangram puzzle. Data collected during this interaction is not part the of our current work. The participants then proceeded to the tangram task. Participants were asked to assemble tangram pieces into a shape using an illustration as a guide. Modeling off of previous HRI work with adult participants [31], one piece of the puzzle was hidden to ensure the participant needed to interact with the instructor to complete the task.

The instructor provided feedback and encouragement to assist the child as they completed the puzzle (see [32] for a description of instructor prompts). None of the instructor prompts were designed to require or promote gesturing. Children were given as much time as necessary to complete the puzzle. For the purpose of this paper, only the tangram task phase with the robot instructor was analyzed.

C. Robot Set-Up

The Misty II was used as the robot instructor for the tangram task due to its child-friendly appearance and interactive features. The robot appeared to be autonomous to the child, although in reality it was being controlled by a human operator in a separate room via a human-centric Wizard of Oz setup that simulated the robot's sensing and reasoning capabilities [33]. While observing the interaction through a one-way mirror, the operator selected from 70 predefined robot behaviors to assist the child. See [32] for a descriptions of the behaviors. To ensure that the behaviors covered all expected scenarios while also limiting the number of controls to allow the operator to quickly find and select the right one, the behaviors and user interface were developed and tested through an iterative development process using a series of test interactions with students and pilot participants. To facilitate rapid and consistent responses by the robot, the operator trained for over 20 hours by practicing in simulated interactions with researchers and pilot participants.

IV. Video Coding

A. Gesture Coding

Session videos during the tangram task phase were coded by research assistants. Task onset was the moment the illustration of the puzzle was presented until the instructor indicated that the participant completed the puzzle.

Based on the gesture categories outlined in [34] and [35], we developed a video coding scheme that categorized hand, arm, and head gestures into three distinct types. [35] utilized a similar scheme on slightly younger children than our sample and also utilized physical objects.

1) **Deictic**: Gestures that direct attention to a specific referent or entity. Based on the literature on child gestures [35], [36], we distinguish between two types of deictic gestures.
    a) **Point**: Gestures that use a finger or hand to point at and direct attention to any object or agent. (E.g, pointing with their finger at the puzzle when looking for the instructor to assess their progress.)
    b) **Show**: Gestures that indicate an object by presenting, holding, or lifting it up. (E.g., picking up a piece and presenting it to the instructor when asking about the piece.)
2) **Conventional**: Gestures that have established social meanings. (E.g., nodding their heads in affirmation or shrugging their shoulders when uncertain what to do.)
3) **Representational**: Gestures to represent objects or actions. (E.g., pretending one's hand is a tangram piece and rotating it in space, or sliding a missing piece into position.)

Three trained research assistants independently coded participant videos using Mangold INTERACT software. Gestures were coded at their onset and were then categorized using the coding scheme above. Twelve out of the 32 videos (37.5%) were double coded by two of the research assistants. The coders then conversed about any disagreements and reached consensus on the final coding file. We analyze total count score (for each gesture type within each instructor block) and gesture rates (gestures/minute) to account for the fact that longer task completion times might provide more time and opportunity to gesture.

V. Results

Our dataset consists of 224 total gestures towards the instructors (robot, human). Children produced 116 gestures with the robot instructor and 108 with the human instructor. We analyzed deictic gestures both separately (show, point) and combined across subtypes. Representational gesture rates appeared too infrequently to analyze individually (n=6), but are included in the total gesture rates.

Total gesture counts were significantly positively correlated with task time (r=0.539, p<.001), such that the longer a child took to complete the task, the more they gestured. Task time also was significantly negatively correlated with age (r=-0.459, p<.001), such that older children completed the task significantly faster than younger children. To account for age differences and to provide a consistent metric across instructor types, the remainder of the analysis uses gesture rates.

TABLE I
Comparing Gesture Rates with Robot vs Human Instructor

	Robot		Human			
	M	SD	M	SD	p	Effect Size
Show	0.218	0.541	0.086	0.191	0.349	0.283
Point	0.166	0.255	0.360	0.443	0.434	-0.199
Deictic	0.383	0.664	0.345	0.484	0.714	0.065
Conventional	0.632	0.747	0.736	0.708	0.494	-0.122
Total	1.031	1.235	1.123	1.041	0.669	-0.076

Note. To calculate the p-values, Show and Point rates use Wilcoxon signed-rank, and Deictic, Conventional, and Total use Student t-test. For the Wilcoxon test, effect size is given by the matched rank biserial correlation. For the Student t-test, effect size is given by Cohen's d.

We compared the rates of each type of gesture and all gesture types combined across the conditions of robot and human instructor. Since show and point gesture rates were not normally distributed based on a Shapiro-Wilk test of normality, we used a Wilcoxon signed-rank test for these gestures. For all of the other gesture rates, we use a paired sample Student's t-test.

The means are higher for show and deictic gestures with the robot instructor and are higher for point, conventional, and total gesture rates with the human instructor (see Table I). In comparing the robot and human instructor conditions, there are no significant differences between rates for show, point, deictic, conventional, and total gestures. However, we do find that when calculating a matched rank biserial correlation for the show and point gesture rates, there is a small to medium effect size of 0.283 and -0.199, respectively. Using a Cohen's d for the other comparisons, we get 0.065 for deictic, -0.122 for conventional, and -0.076 for total. All of these fall below the common standard for a small effect size [37].

A. Age-Related Effects

To look at possible age-related effects on gesture, we calculated a Pearson's correlation coefficient between age and each gesture type in each condition (see Table II and Figure 2). In the robot condition, there was no significant correlation found between age and rates for show, point, deictic, conventional, and total gestures. In the human condition, there was a significant correlation between age and gesture

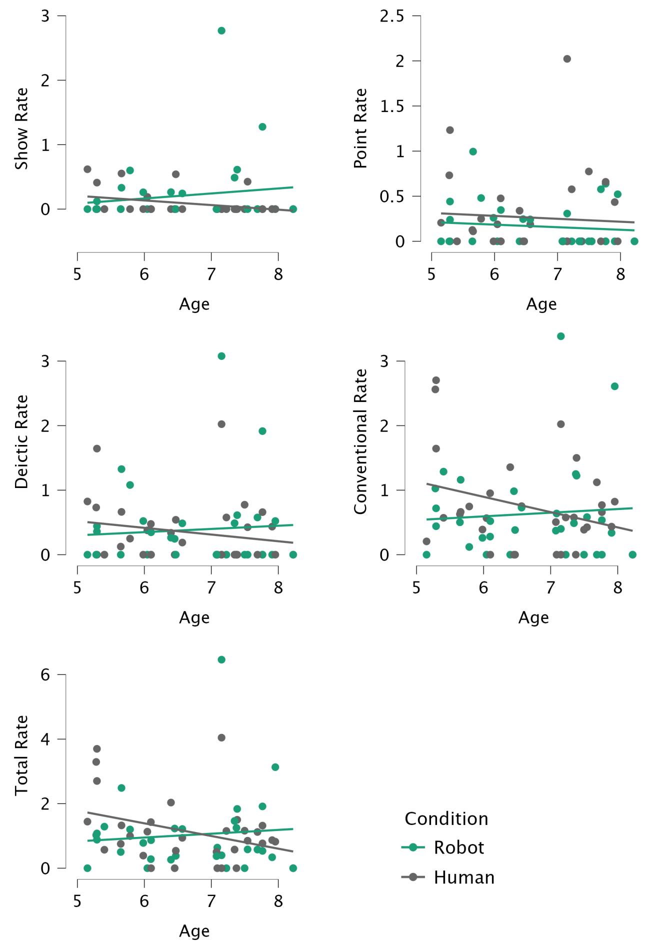

Fig. 2. Scatter plots showing the relationships between age and gesture rates in each condition. Linear regression lines for each condition are displayed.

rates for show, deictic, and total.

None of the correlations between age and gesture rate in the robot condition were significant (all *p*s>.05), but the correlations for the show and total gesture rates in the human condition are significant. Additionally, scatter plots (Fig. 2) seem to show different relationships with age across the two conditions for some of the gesture types. To determine if these correlations are significantly different, we use an online calculator [38] to compare the correlations. The calculator does a test of the difference between two dependent correlations with one variable in common by converting the correlation coefficients to a z-score and then computing the asymptotic covariance [39].

B. Within-subject effects

Since we had some indications that age is related to differences in how much a child uses various gesture types towards a robot versus a human, we expected to find some negative correlations between gesture rates of the same type across the two conditions. This would suggest that if a participant frequently uses a gesture type with one instructor type, then they use it less frequently with the other instructor type. However, this generally was not the case.

Table III shows how rates of each gesture type with one instructor relates to rates of each gesture type with the other instructor. The only significant correlations are positive. Show gesture rate with the robot significant correlates with point (p=0.667), deictic (p=0.560), conventional (p=0.350), and total gesture rates (p=0.486) with the human. Deictic gesture rate with the robot significant correlates with point (p=0.597), deictic (p=0.533), and total gesture rates (p=0.472) with the human. Note that while show and combined deictic gesture rates correlate with many rates with the human, the point gesture rate with the robot does not correlate with any gesture rates with the human. Conventional gesture rates with the robot do not significantly correlate with any gesture rates with the human. Total gesture rate with the robot significant correlates with point (p=0.521), deictic (p=0.443), conventional (p=0.381), and total gesture rates (p=0.486) with the human.

We also conducted a series of repeated measures analysis of covariance (RMANCOVA) to test for differences in how often a child gestures towards a robot versus human instructor when controlling for age. Results for the RMANCOVA on point, show, deictic, conventional, and total gestures rates with age as a covariate are in Table IV. For total gesture rate, there is a significant main effect of instructor type (p=0.023) with a small to medium effect size (ω^2=0.037) and a significant interaction with age (p=0.025) also with a small to medium effect size (ω^2=0.036). For conventional gesture rate, there is not a main effect of instructor type (p=0.059) but a small effect size (ω^2=0.029), and there is no interaction with age (p=0.069) but also a small effect size (ω^2=0.026). For deictic gesture rate, there is not a main effect of instructor type (p=0.188), and there is no interaction with age (p=0.168). The effect sizes are less than small (ω^2=0.006, ω^2=0.007). For point gesture rate, there is not a main effect of instructor type (p=0.849), and there is no interaction with age (p=0.966). The effect sizes are less than small (ω^2=0.000). For show gesture rate, there is not a main effect of instructor type (p=0.264) and less than a small effect size (ω^2=0.005), and there is no interaction with age (p=0.195) but a small effect size (ω^2=0.013).

VI. Discussion

Our study examined children's use of gestures with a robot instructor and a human instructor during a learning task. In contrast to most research in exploring children's gesture

TABLE II

Summary of Correlations with Age

	Robot (r)	Human (r)	Comparison (z-score)
Child vocalization	-0.033	-0.208	1.534
Show rate	0.135	-0.351*	1.820*
Point rate	-0.104	-0.068	-0.149
Deictic rate	0.070	-0.201	0.844
Conventional rate	0.071	-0.314	1.850*
Total rate	0.090	-0.354*	2.369**

* p < .05, ** p < .01, *** p < .001

TABLE III

CORRELATIONS BETWEEN GESTURES RATES WITH HUMAN VS ROBOT

With Robot		With Human	Pearson's r	p
Point Rate	-	Point Rate	0.139	0.448
	-	Show Rate	0.184	0.313
	-	Deictic Rate	0.200	0.273
	-	Conventional Rate	0.161	0.380
	-	Total Gesture Rate	0.197	0.280
Show Rate	-	Point Rate	0.667***	< 0.001
	-	Show Rate	-0.129	0.483
	-	Deictic Rate	0.560***	< 0.001
	-	Conventional Rate	0.350*	0.050
	-	Total Gesture Rate	0.486**	0.005
Deictic Rate	-	Point Rate	0.597***	< 0.001
	-	Show Rate	-0.034	0.852
	-	Deictic Rate	0.533**	0.002
	-	Conventional Rate	0.347	0.052
	-	Total Gesture Rate	0.472**	0.006
Conventional Rate	-	Point Rate	0.343	0.055
	-	Show Rate	-0.103	0.577
	-	Deictic Rate	0.274	0.130
	-	Conventional Rate	0.326	0.069
	-	Total Gesture Rate	0.321	0.074
Total Gesture Rate	-	Point Rate	0.521**	0.002
	-	Show Rate	-0.086	0.639
	-	Deictic Rate	0.443*	0.011
	-	Conventional Rate	0.381*	0.032
	-	Total Gesture Rate	0.441*	0.012

TABLE IV

REPEATED MEASURES ANCOVA WITH AGE AS COVARIATE

Cases	SS	df	MS	F	p	ω^2
Total	3.684	1	3.684	5.731	0.023*	0.037
Total * Age	3.559	1	3.559	5.536	0.025*	0.036
Residuals	19.288	30	0.643			
Conv.	1.278	1	1.278	3.869	0.059	0.029
Conv. * Age	1.175	1	1.175	3.558	0.069	0.026
Residuals	9.907	30	0.330			
Deictic	0.292	1	0.292	1.813	0.188	0.006
Deictic * Age	0.321	1	0.321	1.992	0.168	0.007
Residuals	4.833	30	0.161			
Point	0.004	1	0.004	0.037	0.849	0.000
Point * Age	$< .001$	1	$< .001$	0.002	0.966	0.000
Residuals	3.562	30	0.119			
Show	0.225	1	0.225	1.296	0.264	0.005
Show * Age	0.305	1	0.305	1.753	0.195	0.013
Residuals	5.212	30	0.174			

production that use narratives to promote gesture production (e.g., [5]) or requires children to gesture to communicate (e.g., [40]), our task was not explicitly designed to elicit gestures. Rather, children's gestures were spontaneous and not prompted by the instructor. The spontaneous emergence of children's gestures within a task that did not actively encourage gesturing indicates that gesture use may be significantly prevalent in interactions design to elicit or support gestures, highlighting the importance of systematically investigating children's gestures in various interaction settings.

We coded for three types of gestures: deictic, conventional, and representational. We further broke deictic gesture down into two sub-types: show and point. We found that children rarely used representational gestures, and therefore we did not include these gestures in our final analysis. Since the number of gestures presented by a participant corresponded to how long the participant took to complete the tangram task, our analysis uses gesture rates (gestures/minute). When comparing the rates of all gesture types, our results show no significant differences between the rates at which children gesture toward the robot and human instructors.

However, our results suggest that age influences how frequently children gesture with the different instructor types. When we compare gesture rates towards each instructor type and control for age as a covariate, we see a significant difference in the total gesture rates. Similarly, we also find a significant difference with total gesture rates when comparing the correlation between age and the total gesture rates for robot and human. Looking at our scatter plots (see Figure 2), we see that younger children in our sample tended to use more gestures with the human instructor than the robot, but older children were more likely to use fewer gestures with the human instructor compared to the robot instructor.

Some of these differences may be related to perceptions of the robot, relative to the human. Prior research has shown that adults have different views of cognitive abilities of humans versus robots [41]–[43]. For children in our sample, the younger children may perceive the robot to have insufficient ability to help them. Older children in our sample may be more willing to engage with the robot because they have some belief that it can help. In examining the video for participant A, who was 5-years old, we saw her turn to the researcher in the room for help when working with the robot instructor. Conversely, participant B, who was 7-years old, engaged with the robot, asking it questions while using many gestures.

It probably can be assumed that the older children in our sample perceived the human instructor to be capable of helping. However, they may have been less willing to engage with the human instructor. Older children may be more independent. Another reason could be that as children get older, their social perceptions of the robot may change. Prior research has shown that university students may seek help from a human instructor less frequently than a robot instructor because the robot is perceived to have a lower social status than the human [44].

The age-related effects could also be related to how difficult the task is for children at various ages, as is evident in the strong negative correlation between the participants age and time to complete the task. Children who found the task less challenging had fewer reasons to engage with the robot during the task. See Section VI-C for further discussion on how gesturing may be related to children seeking help or feedback.

A. Deictic Gestures

In the research on deictic gestures, there has been a predominant focus on point gestures. Show gestures, which direct the focus of attention on an object by holding it, are distinct from pointing and are developed before pointing

[45]. Whereas a pointing gesture requires the receiver of the gesture to redirect their attention, a show gesture is simpler in that the object is presented in the recipient's field of vision. It can then be expected that show gestures could be more commonly performed with a robot since there may be a belief that the object needs to be held up directly in front of the robot's camera in order for it to see it.

In our study, we found a complex relationship between types of deictic gestures. When combining the types of deictic gestures, our results showed that the rate of deictic gestures with the robot correlates with the rate of deictic gestures with the human. However, through employing a coding scheme to separately examine the rates for the show and point gestures, we found that they did not correlate across conditions. This means that a child's rate of using a show (or point) gesture with a human instructor did not correspond with how they used show (or point) gestures with a robot. Since the overall rates of deictic gestures did correlate, this result seems to indicate that a child may have been more inclined to use one type of deictic gesture with one instructor type and shift to the other type with the other instructor. In fact, we saw that point gesture rates with a human instructor correlated with the show gesture rates with the robot, suggesting that a child may have used a show gesture with a robot when they used a point gesture with a human. The pattern is exemplified by participant B, who had a point rate of 2.02 gestures/min. and show rate of 0.00 gestures/min. with the human instructor and a point rate of 0.31 gestures/min. and show rate of 2.77 gestures/min. with the robot. Future research should continue to examine both types of deictic gestures in child-robot interactions to fully capture children's use of gesture in this context.

Our results also showed there was a significant difference in the correlations between age and show gestures across instructor types. The older the child was, the less frequently they tended to use a show gesture with a human instructor. The significant difference in the correlations means that age is more related to show gestures with a human than with a robot. Age at least partially explains the differences in how a child used a show gesture with one instructor but not the other. However, in our sample, age did not show any relation to point gesture rates, and thus it remains unknown what factors contributed to a child's different uses of pointing gestures. To gain further understanding of the child's behavior, it will be important to examine how individual differences relate to a child's behavior with a human versus a robot [46].

Our findings indicated that show and point gestures had different relationships in a learning setting and did not necessarily follow the same patterns of deictic gestures overall. This result emphasizes distinguishing between the types of deictic gestures is needed to better understand children's use of deictic gestures. Our coding scheme made this distinction. For future research to lead to a better understanding of the factors contributing to a child's use of deictic gestures with a robot, coding and analysis needs to use similar coding schemes that distinguish between the types of deictic gestures.

B. Conventional Gestures

Conventional gestures are less frequently studied in the context of human-robot interactions, but our study found that children regularly exhibited spontaneous use of conventional gestures like head nods and shrugs. Though there was no significant difference in rates of conventional gesture across instructor type, we cannot conclude that children used these gestures in a similar fashion with a robot as with a human since the gesture rates did not significantly correlate ($p=0.069$). Additionally, we see a small effect size in our RMANCOVA for conventional gesture rates. The small effect size and low significance values suggest that we might see a significant trend with a larger participant pool.

However, we found differences in conventional gesture rates when accounting for age. For the correlations between age and conventional gesture rates, we found a significant difference in the correlations for the robot and human conditions. This means that age was more related to conventional gesture rates with a human than with a robot. Additionally, the conventional gesture rate with a robot did not significantly correlate with the rates of any of the gesture types with the human. With this data, it remains unclear what factors are contributing to how often a child gestures with a robot.

In reviewing the videos for a few participants, we were able to speculate that one of the factors contributing to the use of conventional gestures was whether the child desired to use the gesture in substitution for a vocalization. Children nodded or shook their heads in response to the instructor's question, and the child does not accompany the gesture with any vocalization. For example, participant C shook his head when the human instructor asked if he had all of the pieces, but he vocalized a response to the same question when the robot instructor asked it. Conventional gestures are interesting in that their established meanings allow them to be used in replace of language [47], but our current analysis did not help us predict whether or not the child used a conventional gesture. Future work may need to consider how a child's individual characteristics relate to their gesturing behavior, as recent results have shown an influence of these characteristics on other social behaviors such as gaze patterns and social referencing [46].

Other conventional gestures were used to express confusion, uncertainty, or contemplation. For example, participants D and E shrug when unsure if they were correctly following the robot's instruction on how to move a piece. Participant A waved her arms over her head when she was unsure and waited for the instructor to give her feedback. Participant B moved her arms to her sides or hips when appearing to be uncertain and brought her hands to her face when thinking. She presented these gestures with both instructors, though more frequently with the robot. While most HRI research on conventional or symbolic gestures focus on head nods or shakes, these examples show there are other conventional gestures that children naturally use to communicate with the instructor. Importantly, they were communicating information about their mental state and whether they needed the instructor's feedback. These are critical cues for a social robot to understand when assisting a child.

C. Implications for HRI Design and Applications

This is the first study in HRI to examine how children use different types of gestures towards a robot and human instructor. Our results provide evidence that children naturally and spontaneously use show, point, and conventional gestures when working with a social robot. Since children's gestures were not accompanied by speech, the only way for the robot to understand what the child was communicating is for the robot to understand the child's gestures. HRI researchers and designers cannot solely rely on past research on children's gestures, since most of that research explores how children gesture with a human partner. Our results suggest that how children gesture towards humans may not necessarily match how they will gesture with a robot. Thus it is critical we have more research addressing children's use of gestures with robots, and our research serves this goal by identifying the types and frequencies of gestures that children use.

In closely examining the videos of gestures produced by children in our study, we noticed that a number of the gestures were used to express confusion or to seek help. Interestingly, many of these examples also did not have co-speech. The child was expecting the instructor to be able to recognize and interpret their nonverbal behavior. For example, when participant D was learning with the robot instructor, she paused to point at the puzzle when she needed to direct the robot's attention to it because she needed help. She did not say anything, only pointed at the puzzle while looking at the robot. Prior work has shown the value of interpreting these types of gaze patterns to recognize when adult users need or want help from a robot [31], [48], and we now have reason to believe that in addition to gaze patterns, deictic and conventional gestures would have a similar value with child populations.

Gestures play a key role in supporting synchronization and mutual attention, essential features for the development of rapport [49]. As such, gestures have been shown to be related to rapport in human-human interactions [50], human-virtual agent interactions [51], and human-robot interactions [4], [52]. While prior research with virtual agents and robots focused on gestures performed by the artificial agent, it is reasonable to conclude that recognizing a child's head nod would support the child's sense of feeling in sync with the robot and recognizing their pointing to the puzzle facilitates mutual attention.

D. Limitations

A human instructor has more expressive capabilities than a social robot, which could incline a child to use more social behaviors with a human as opposed to a robot. To control for this, the human instructor in our study was trained to limit their responses to the behaviors that the robot was programmed to perform. With the human instructor not using voice inflections, large gestures and few facial expressions, children may have found the human instructor to be less engaging.

We recognize that conventional gestures are highly coupled with culture, in both their semantics and the frequency with which they are used. For example, Japanese speakers produced more nodding during a robot's speaking than English speakers in a human-robot interaction, but no differences in the frequency of head gestures was found in human-human interactions [53]. Future work could examine how children across the world differ in their rate, size, and preferred mode of gesturing to robots.

VII. Conclusion

We provide evidence that children spontaneously use a variety of gesture types with a robot. We also find that the children's rate of gesturing with a robot does not necessarily match their rates of gesturing with a human instructor. Age shows to be a factor in how much a child gestures with a human instructor, but age is not associated with gesturing rates with a robot instructor. We also highlight the need for a coding scheme that distinguishes between gesture types. Through the application of our gesture coding scheme we are able to identify differences in how children use show gestures versus point gestures. Finally, given the lack of existing research in children's gestures in HRI, our work lays a foundation for further investigation into how and why children gesture with robots in a learning environment.

References

[1] M. A. Novack and S. Goldin-Meadow, "Gesture as representational action: A paper about function," *Psychonomic Bulletin & Review*, vol. 24, pp. 652–665, 2017.

[2] M. A. Novack, S. Goldin-Meadow, and A. L. Woodward, "Learning from gesture: How early does it happen?" *Cognition*, vol. 142, pp. 138–147, 2015.

[3] D. McNeill, E. T. Levy, and S. D. Duncan, "Gesture in discourse," *The handbook of discourse analysis*, pp. 262–289, 2015.

[4] J. R. Wilson, N. Y. Lee, A. Saechao, S. Hershenson, M. Scheutz, and L. Tickle-Degnen, "Hand gestures and verbal acknowledgments improve human-robot rapport," in *Social Robotics: 9th Intl. Conf.* Springer, 2017, pp. 334–344.

[5] J.-M. Colletta, C. Pellenq, and M. Guidetti, "Age-related changes in co-speech gesture and narrative: Evidence from french children and adults," *Speech Communication*, vol. 52, no. 6, pp. 565–576, 2010.

[6] M. A. Evans and K. H. Rubin, "Hand gestures as a communicative mode in school-aged children," *The J. of Genetic Psychology*, vol. 135, no. 2, pp. 189–196, 1979.

[7] E. L. Congdon, E. M. Wakefield, M. A. Novack, N. Hemani-Lopez, and S. Goldin-Meadow, "Learners' spontaneous gesture before a math lesson predicts the efficacy of seeing versus doing gesture during the lesson," *Cognitive Science*, vol. 48, no. 7, p. e13479, 2024.

[8] E. L. Congdon, M. A. Novack, and E. M. Wakefield, "Exploring individual differences: A case for measuring children's spontaneous gesture production as a predictor of learning from gesture instruction," *Topics in Cognitive Science*, 2024.

[9] M. W. Alibali and S. Goldin-Meadow, "Modeling learning using evidence from speech and gesture," in *Proc. of the Fifteenth Annual Conf. of the Cognitive Science Society*. Erlbaum Hillsdale, NJ, 1993, pp. 203–208.

[10] E. A. Konijn, M. Smakman, and R. Berghe, "Use of robots in education," in *The international encyclopedia of media psychology*. Wiley, 2020, pp. 1–8.

[11] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, "Social robots for education: A review," *Science Robotics*, vol. 3, no. 21, p. eaat5954, 2018.

[12] C.-M. Huang and B. Mutlu, "Modeling and evaluating narrative gestures for humanlike robots." in *Robotics: Science and Systems*, vol. 2. Citeseer, 2013.

[13] H. Kose-Bagci, E. Ferrari, K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, "Effects of embodiment and gestures on social interaction in drumming games with a humanoid robot," *Advanced Robotics*, vol. 23, no. 14, pp. 1951–1996, 2009.

[14] J. de Wit, A. Brandse, E. Krahmer, and P. Vogt, "Varied human-like gestures for social robots: Investigating the effects on children's engagement and language learning," in *Proc. of the 2020 ACM/IEEE Intl. Conf. on Human-Robot Interaction*, 2020, pp. 359–367.

[15] T. Komatsubara, M. Shiomi, T. Kanda, and H. Ishiguro, "Can using pointing gestures encourage children to ask questions?" *Intl. J. of Social Robotics*, vol. 10, pp. 387–399, 2018.

[16] M. de Haas, P. Vogt, R. van den Berghe, P. Leseman, O. Oudgenoeg-Paz, B. Willemsen, J. de Wit, and E. Krahmer, "Engagement in longitudinal child-robot language learning interactions: Disentangling robot and task engagement," *Intl. J. of Child-Computer Interaction*, vol. 33, p. 100501, 2022.

[17] H. Liu and L. Wang, "Gesture recognition for human-robot collaboration: A review," *Intl. J. of Industrial Ergonomics*, vol. 68, pp. 355–367, 2018.

[18] L. Fiorini, F. G. C. Loizzo, A. Sorrentino, J. Kim, E. Rovini, A. Di Nuovo, and F. Cavallo, "Daily gesture recognition during human-robot interaction combining vision and wearable systems," *IEEE Sensors J.*, vol. 21, no. 20, pp. 23 568–23 577, 2021.

[19] J. Qi, L. Ma, Z. Cui, and Y. Yu, "Computer vision-based hand gesture recognition for human-robot interaction: a review," *Complex & Intelligent Systems*, vol. 10, no. 1, pp. 1581–1606, 2024.

[20] T. Williams, S. Acharya, S. Schreitter, and M. Scheutz, "Situated open world reference resolution for human-robot dialogue," in *2016 11th ACM/IEEE Intl. Conf. on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 311–318.

[21] T. Williams, M. Bussing, S. Cabrol, E. Boyle, and N. Tran, "Mixed reality deictic gesture for multi-modal robot communication," in *2019 14th ACM/IEEE Intl. Conf. on Human-Robot Interaction (HRI)*. IEEE, 2019, pp. 191–201.

[22] C. Matuszek, L. Bo, L. Zettlemoyer, and D. Fox, "Learning from unscripted deictic gesture and language for human-robot interactions," in *Proc. of the AAAI Conf. on Artificial Intelligence*, vol. 28, no. 1, 2014.

[23] N. Otero, C. L. Nehaniv, D. S. Syrdal, and K. Dautenhahn, "Naturally occurring gestures in a human–robot teaching scenario," *Interaction Studies*, vol. 9, no. 3, pp. 519–550, 2008.

[24] D. C. Silpani, K. Suematsu, and K. Yoshida, "A feasibility study on hand gesture recognition in natural conversation," in *2021 5th IEEE Intl. Conf.on Cybernetics (CYBCONF)*. IEEE, 2021, pp. 085–090.

[25] C. Di Dio, F. Manzi, G. Peretti, A. Cangelosi, P. L. Harris, D. Massaro, and A. Marchetti, "Shall i trust you? from child–robot interaction to trusting relationships," *Frontiers in psychology*, vol. 11, p. 469, 2020.

[26] X. Li and W. Q. Yow, "Younger, not older, children trust an inaccurate human informant more than an inaccurate robot informant," *Child Development*, vol. 95, no. 3, pp. 988–1000, 2024.

[27] P. Baxter, E. Ashurst, R. Read, J. Kennedy, and T. Belpaeme, "Robot education peers in a situated primary school study: Personalisation promotes child learning," *PloS one*, vol. 12, no. 5, p. e0178126, 2017.

[28] C. L. Van Straten, J. Peter, and R. Kühne, "Child–robot relationship formation: A narrative review of empirical research," *Intl. J. of Social Robotics*, vol. 12, no. 2, pp. 325–344, 2020.

[29] B. N. Verdine, C. M. Irwin, R. M. Golinkoff, and K. Hirsh-Pasek, "Contributions of executive function and spatial skills to preschool mathematics achievement," *Journal of experimental child psychology*, vol. 126, pp. 37–51, 2014.

[30] N. M. Siew, C. L. Chong, and M. R. Abdullah, "Facilitating students'geometric thinking through van hiele's phase-based learning using tangram," *J. of Social Sciences*, vol. 9, no. 3, p. 101, 2013.

[31] J. R. Wilson, P. T. Aung, and I. Boucher, "When to help? a multimodal architecture for recognizing when a user needs help from a social robot," in *International Conference on Social Robotics*. Cham: Springer Nature Switzerland, 2022, pp. 253–266.

[32] Y. Yang, A. Langer, L. Howard, P. J. Marshall, and J. R. Wilson, "Towards an ontology for generating behaviors for socially assistive robots helping young children," in *Proc. of the AAAI Symposium Series*, vol. 2, no. 1, 2023, pp. 213–218.

[33] L. D. Riek, "Wizard of oz studies in hri: a systematic review and new reporting guidelines," *J. of Human-Robot Interaction*, vol. 1, no. 1, pp. 119–136, 2012.

[34] M. David, "Hand and mind: What gestures reveal about thought," *Univ. of Chicago press.*, 1992.

[35] M. A. Novack, M. Standley, M. Bang, K. Washinawatok, D. Medin, and S. Waxman, "Hands on: Nonverbal communication in native and non-native american parent–child dyads during informal learning." *Developmental Psychology*, vol. 58, no. 1, p. 32, 2022.

[36] B. Choi, R. Wei, and M. L. Rowe, "Show, give, and point gestures across infancy differentially predict language development." *Developmental Psychology*, vol. 57, no. 6, p. 851, 2021.

[37] J. Cohen, "A power primer," 2016.

[38] I. A. Lee and K. J. Preacher, "Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]," 2013, Available from http://quantpsy.org.

[39] J. H. Steiger, "Tests for comparing elements of a correlation matrix." *Psychological bulletin*, vol. 87, no. 2, p. 245, 1980.

[40] T. Behne, M. Carpenter, and M. Tomasello, "Young children create iconic gestures to inform others." *Developmental psychology*, vol. 50, no. 8, p. 2049, 2014.

[41] D. T. Levin, S. S. Killingsworth, and M. M. Saylor, "Concepts about the capabilities of computers and robots: A test of the scope of adults' theory of mind," in *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, 2008, pp. 57–64.

[42] D. T. Levin, M. M. Saylor, and S. D. Lynn, "Distinguishing first-line defaults and second-line conceptualization in reasoning about humans, robots, and computers," *International Journal of Human-Computer Studies*, vol. 70, no. 8, pp. 527–534, 2012.

[43] R. Q. Stafford, B. A. MacDonald, C. Jayawardena, D. M. Wegner, and E. Broadbent, "Does the robot have a mind? mind perception and attitudes towards robots predict use of an eldercare robot," *Intl. J. of Social Robotics*, vol. 6, pp. 17–32, 2014.

[44] I. Howley, T. Kanda, K. Hayashi, and C. Rosé, "Effects of social presence and social role on help-seeking and learning," in *Proc. of the ACM/IEEE Intl. Conf. on Human-Robot Interaction*, 2014, pp. 415–422.

[45] T. Cameron-Faulkner, A. Theakston, E. Lieven, and M. Tomasello, "The relationship between infant holdout and gives, and pointing," *Infancy*, vol. 20, no. 5, pp. 576–586, 2015.

[46] A. Langer, J. R. Wilson, L. Howard, and P. Marshall, "The influence of child characteristics on children's behavior towards a social robot versus a human," under review.

[47] M. Guidetti, "The emergence of pragmatics: forms and functions of conventional gestures in young french children," *First Language*, vol. 22, no. 3, pp. 265–285, 2002.

[48] U. Kurylo and J. R. Wilson, "Using human eye gaze patterns as indicators of need for assistance from a socially assistive robot," in *Social Robotics: 11th International Conference*. Cham: Springer International Publishing, 2019, pp. 200–210.

[49] L. Tickle-Degnen and R. Rosenthal, "The nature of rapport and its nonverbal correlates," *Psychological inquiry*, vol. 1, no. 4, pp. 285–293, 1990.

[50] J. E. Grahe and F. J. Bernieri, "The importance of nonverbal cues in judging rapport," *J. of Nonverbal behavior*, vol. 23, pp. 253–269, 1999.

[51] J. Gratch, A. Okhmatovskaia, F. Lamothe, S. Marsella, M. Morales, R. J. van der Werf, and L.-P. Morency, "Virtual rapport," in *Intelligent Virtual Agents: 6th International Conference*. Springer, 2006, pp. 14–27.

[52] L. D. Riek, P. C. Paul, and P. Robinson, "When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry," *J. on Multimodal User Interfaces*, vol. 3, pp. 99–108, 2010.

[53] T. Mori, K. Jokinen, and Y. Den, "Analysis of body behaviours in human-human and human-robot interactions," in *Proceedings of LREC2020 Workshop" People in language, vision and the mind"(ONION2020)*, 2020, pp. 7–14.