How to Train Your DragonBot: Socially Assistive Robots for Teaching Children About Nutrition Through Play

Elaine Short[1,2], Katelyn Swift-Spong[2], Jillian Greczek[2], Aditi Ramachandran[3], Alexandru Litoiu[3], Elena Corina Grigore[3], David Feil-Seifer[4], Samuel Shuster[2], Jin Joo Lee[5], Shaobo Huang[2], Svetlana Levonisova[2], Sarah Litz[6], Jamy Li[7], Gisele Ragusa[2], Donna Spruijt-Metz[2], Maja Matarić[2], and Brian Scassellati[3]

Abstract— **This paper describes an extended (6-session) interaction between an ethnically and geographically diverse group of 26 first-grade children and the DragonBot robot in the context of learning about healthy food choices. We find that children demonstrate a high level of enjoyment when interacting with the robot, and a statistically significant increase in engagement with the system over the duration of the interaction. We also find evidence of relationship-building between the child and robot, and encouraging trends towards child learning. These results are promising for the use of socially assistive robotic technologies for long-term one-on-one educational interventions for younger children.**

I. Introduction

Children love imaginary play, and such play can make learning more engaging and effective [1], [2]. There is also much need for individualized support in a classroom, allowing each child to progress at their own speed using the learning strategies most appropriate for that child. Child-friendly social robots have the potential to provide this individualized support through imaginary play, allowing children to engage with their lessons in a tangible and active way.

We take a Socially Assistive Robotics (SAR) approach to teaching nutrition to 1st-grade children, with the goal of promoting positive habits and behavior change through human-robot interaction (HRI). In order to move towards the kind of long-term deployment that would be necessary to see major gains in learning (consider that a single topic might be covered in a school setting for many weeks), we first evaluate the feasibility of this approach, including measuring children's engagement with the SAR system over time and evaluating whether it has the potential to facilitate nutrition information learning. We present a semi-autonomous SAR system, and a three-week biweekly evaluation study with a diverse group of 26 5-8 year old children.

*This work was supported by the National Science Foundation (Expeditions in Computing IIS-1139148, CNS-0709296, REU Supplements, and GRFP).

[1]Corresponding author: elaine.g.short@usc.edu
[2]University of Southern California
[3]Yale University
[4]University of Nevada, Reno
[5]Massachusetts Institute of Technology
[6]State University of New York
[7]Stanford University

II. Background

Childhood obesity has tripled in the United States over the past 4 decades [3]. Obesity among children and adolescents has been shown not only to lead to increase risk of being overweight in adulthood[4], but also diseases later in life, including high cholesterol and triglycerides, hypertension, and type 2 diabetes [5]. Educating children about healthy food and beverage choices, and motivating them to make healthier choices, can help to lower rates of obesity [6]. Technological interventions in particular lend themselves to the broad replication and personalization that will be necessary to combat what has proven to be a challenging problem [7].

Several technology-based nutrition interventions for children have been developed, using smartphones, computers, and video games (see [8] or [9] for a review). While these interventions make use of technologies that are widely available, there is some evidence that HRI systems promote learning more than screen-based technologies. Leyzberg et al. showed that the physical presence of a robot can increase cognitive learning gains [10], while in children, Movellan et al. [11] demonstrated that a social robot could be used to teach young children new words, despite other results suggesting that young children do not learn language from pre-recorded human speech[12]. Other work in HRI focuses on the use of robotic teachers for english language learning [13], [14] and teaching children to play chess [15]. In adults, Kidd and Breazeal [16] show that SAR systems have the potential to improve eating and exercise beyond a paper- or computer-based intervention.

Our work aims to use SAR to leverage children's excitement about both pretend play and technology, to provide an affordable, accessible, and personalized means of delivering nutrition education and coaching. In this paper we focus on teaching first-grade children (primarily 6-7 years old, although several participants were 5 or 8 years old at the time of the study). Children this age do not have significant control over their food choices, allowing us to instill healthy habits before unhealthy ones become ingrained.

III. Methodology

In the sections that follow, we describe the design of a SAR-based nutrition education intervention, as well as the study used to test the effects of the initial version of the intervention. Our study addresses five research questions, as follows: *Q1:* Do children enjoy interacting with the SAR system? *Q2:* Are children able to maintain engagement with a SAR system over time? *Q3:* Are children able to build a relationship with the SAR system over time? *Q4:* What is the impact of the SAR system on child learning of nutrition information? *Q5:* What is the relationship between the child's temperament and their interaction with the SAR system?

A. *Intervention Design*

While most nutrition education interventions are 12-16 weeks long, most SAR systems are used in much shorter-term interactions (often as little as one session). In order to move towards a longer intervention, we developed a 3-week long, twice-weekly intervention for first grade children, with a within-subjects design. Each of the six intervention sessions consists of an approximately 5- to 10-minute long one-on-one interaction between the child and the robot. We used a Wizard-of-Oz interaction and monitoring design, with a teleoperator providing dialogue selection and some perceptional capabilities, with pre-scripted dialogue behaviors (including both speech and movement), and autonomous control of the overall interaction flow.

Each session consisted of two parts, introduction and food selection game, as described in section III-D. In the first session of each week, the robot acted as an expert, giving feedback on food choices one-by-one, while in the second session of the week the child and robot collaborated toward making healthy choices together.

B. *Robot and Experimental Setup*

The intervention centered around an interaction with the DragonBot robot [17], a dragon-like squash-and-stretch robot with five degrees of freedom (see Figure 2, center-left). The robot is covered with a plush skin designed in collaboration with an expert puppeteer. The skin includes posable arms and tail, as well as removable wings in four sizes, allowing the robot's wings to "grow" from session to session. The robot is approximately 18 inches tall at its full height and can be seen in Figure 2.

The interaction was conducted in a small parent-teacher conference room, allowing children to interact individually with the robot. The robot was set up on a table facing the child, with realistic artificial foods arranged around it; the arrangement of foods was kept constant across participants. In order to provide the richest possible dataset for future analysis, a Microsoft Kinect sensor, an HD camera, and USB camera are arranged as seen in Figure 1, in addition to the two laptops, the DragonBot base station (providing power to the robot), and set of speakers necessary to run the robot. The intervention setup can be seen in Figure 2 (the cameras and Kinect are out of the frame).

C. *Interaction Structure and Progression*

The verbal interaction between the child and robot used pre-recorded speech following a script that was written with

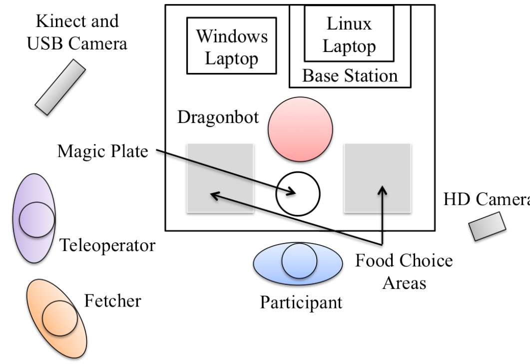

Fig. 1. Diagram of the intervention setup

Fig. 2. Intervention room setup

the assistance of a screenwriter experienced in children's television, and follows the story of Chili the DragonBot through several weeks of training and preparation for the big upcoming "dragon race". We employed child-centric storytelling techniques such as character development and backstory to create a richer interaction. We also increase the difficulty of the task over the three-week intervention, challenging the child within her or his zone of proximal development, using a socio-constructivist approach [18], [19] to integrate the increasing difficulty and social interaction, maximizing the child's learning potential.

Over the course of a six-session, three-week experiment, we covered three nutrition topics, one per week: packing a lunch box (choosing whole grains and non-sugary drinks), choosing after-school snacks (avoiding nutritionally bankrupt junk food), and building balanced meals (choosing whole grain breakfast cereals and colorful vegetables at dinner). Each topic had two sessions devoted to it; each session built upon the previous one, using a gradual increase in challenge of content. In the first session, the robot served as an expert, sampling foods offered to it one at a time and providing nutritional information as feedback. In the second session, the robot behaved in a more cooperative role, in order to provide feedback in a more challenging game where the child chooses several foods at one time. We refer to the first of these as the expert sessions (ES) and the second as the cooperative sessions (CS).

D. Interaction Segments

Before the beginning of the interaction, the participating child was brought from the classroom by an adult experimenter (not the teleoperator). This person, the "fetcher", remained in the room, sitting behind the child, instructed to answer as simply as possible any questions asked by the child, and to make sure that the child did not damage the robot. The child was invited to sit down, and told to wake up the robot. Once the child said something to the robot, the robot woke up and began the first segment, the introduction. The introduction included a welcome speech, relationship-building through small talk, and finally backstory and character progression. During the CS, the second day on a given nutritional topic, there was a brief review of the nutritional curriculum from the previous session. The interaction then moved into the food choice segment. For the ES, there were two rounds of the food choice game, where the child was asked to choose a food, the robot "tasted" the food, and then provided feedback. In the CS, the child was asked to select a food items until they chose a healthy item, receiving progressively more specific advice about how to improve the selections, based on an evaluation external to both robot and child (in this case, a "magic plate"), and following the theory of graded cueing, an occupational therapy technique [20]. For both session types, some vocabulary used in the feedback was explained with a backstory-type dialogue item.

E. Data Collection and Associated Measures

We recorded several types of electronic data during the intervention in order to build a rich dataset, including teleoperator selections, pre- and post-questionnaires (modified to be administered orally by an experimenter), audio- and video- data, and Kinect pose data. We also collected data from four questionnaires widely used either in the robotics or child development research fields as measures of child personality, child evaluation, and child interaction. The *Child Behavior Questionnaire* [21] (CBQ-S; Cronbach's α= 0.87) contains a 4-point Likert-type scale and requires parents to rate various aspects of their child's behavior and personality (affect). It contains three subscales and provides an efficient child temperament measure for school aged children (ages 5-12). The CBQ-S subscales are surgency (positive affect), effortful control (self-regulation) and negative affect. The following three were administered directly to the children, in an interview-type format to accommodate the children's limited cognitive and developmental abilities and to avoid biases associated with diversity in child reading abilities [22]. The *Perceived Value Questionnaire* is adapted from an evaluative questionnaire by Lombard [23] and also used in SAR related research by Kidd and Breazeal [24] (Cronbach's α = 0.95). This questionnaire required child participants to rate their interaction with the robot using an 8-point Likert-type scale. The "utility" and "value" subscales of this questionnaire were administered after the first interaction that the child had with the robot, and then again after the final interaction with the robot at the culmination of the intervention. The *Social Presence Questionnaire* was used to quantify the effectiveness of the robot's social capabilities (or social presence). The social presence of the robot was measured by an 8-point Likert-type scale using questionnaire items established from Jung and Lee [25], (Cronbach's α = 0.82). We administered this questionnaire after the first intervention session and at the end of the intervention. Finally, the *Adapted Companion Animal Bonding Scale* [26] asks the child to rate the various features of the robot including whether the robot is bad/good, loving/not loving, cuddly/not cuddly, and warm/cold. This questionnaire was administered twice, first before the children interacted with the robot (but after a brief introduction to the robot), and again at the culmination of the intervention.

F. Hypotheses

Six research hypotheses were developed that are associated with the study's research questions:

H1: Participants will have a positive reaction to the SAR system, that will increase over time.

H2: Children will be more engaged with the robot over time, as measured by a decrease in their response time to the robot's verbal questions (a well-established proxy for engagement in the child development literature [27]).

H3: Children will use more complex speech with the robot over time, as measured by mean length of utterance (MLU) and a qualitative categorization of their utterances.

H4: Children's knowledge of, and comfort with, nutritional information will increase, decreasing their time to make a choice of food when prompted by the robot.

H5: Children's performance on the nutrition task will improve over time, as measured by a choice indicating ratio.

H6: Children with a positive affect and higher self-regulatory ability will have greater interaction with the robot.

G. Study Population

As we hope to create systems that can be deployed with children across a wide range of economic and social backgrounds, we collected data from two highly diverse sites within the United States: a west coast site in an urban center and an east coast site that drew from primarily suburban households. We treated these samples as one cohort in our analysis to highlight commonalities relevant to a more robust long-term deployment. There were 26 participants in the study with age range of 5-8 (twenty-two 6-7-year-olds; two 5-year-olds; two 8-year-olds). Seventeen of the children were female (65%) and the remaining nine were male (35%). In terms of ethnicity, the sample was diverse and representative of the areas in which the participants reside. The largest ethnic group represented was children of Hispanic descent (69%), 19% were children of African American descent and the remaining 6% were European American or had mixed ethnicities. Approximately 62% of the study participants reside in the western United States in a large urban city and the remaining study participants (38%) came from the eastern most region of the U.S. The participating children's parents' ages ranged between 20-49, with 25% of the parents in the 20-29 year age range, 44% between 30-39 years of

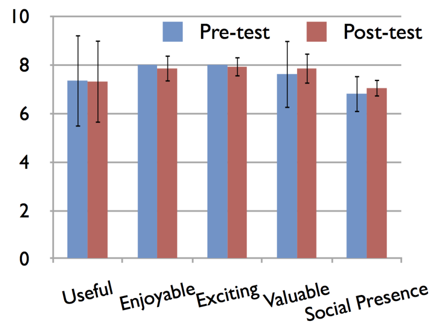

Fig. 3. Child evaluation of the robot in several categories

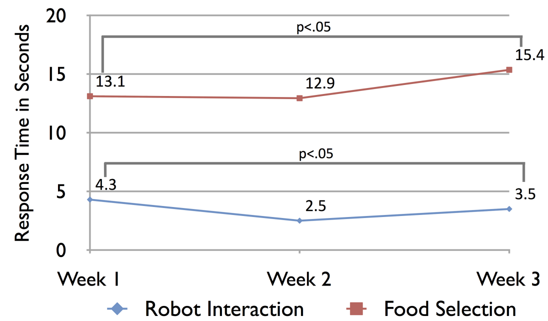

Fig. 4. Child response times to robot conversational queries and food selection prompts

age, and 31% between 40-49 years of age. With regard to parental education, approximately 19% of the children's parents did not graduate from high school, 19% were high school graduates, 43% had completed some college, 19% of the parents had bachelor's degrees, and the remaining 6% had completed graduate or professional education.

IV. Results

A. Evaluation and Perception of the Robot

In general, the participants in our study had positive perceptions and reactions to the socially assistive robot before, during, and and after the SAR intervention. The children's perception of the robot was high ($M = 7.58, SD = .76$, 8-pt. scale) pre-intervention and remained relatively constant through the culmination of the intervention ($M = 7.45, SD = .80$). Specific to the measures of robot evaluation, the children perceived the robot as useful ($M_{PRE} = 7.35, SD = 1.9; M_{POST} = 7.31, SD = 1.7$) and enjoyable ($M_{PRE} = 8, SD = 0; M_{POST} = 7.85, SD = .54$). They also rated it as exciting ($M_{PRE} = 8, SD = 0; M_{POST} = 7.92, SD = .40$), valuable ($M_{PRE} = 7.62, SD = 1.4; M_{POST} = 7.85, SD = .63$), as having strong social presence ($M_{PRE} = 6.81, SD = .76; M_{POST} = 7.04, SD = .43$), and as attractive ($M_{PRE} = 6.65, SD = 1.7; M_{POST} = 7.89, SD = .35$). We did not find significant differences between the pre-and post-intervention ratings, as seen in Figure 3, likely due to the extremely high positive evaluation. Thus the first part of *H1*, the positive perception of the robot, is supported, but not that there is an increase in this perception over the duration of the intervention.

B. Child-Robot Interaction

To determine the level of engagement between the robot and the child during the intervention, we calculated the children's mean response time (in seconds) when prompted by the robot's verbal questions. Although we found that child response times ebbed and flowed across intervention sessions, the comparative results of these calculations revealed that the mean child response time decreased from the first day of the intervention to the end of the intervention. The mean response time across children during the first intervention was 4.3 seconds and the mean response time was 3.5 for the last session, indicating a .8 second mean decrease in response time. A paired t-test was performed to ascertain whether there was statistically significant response efficiency between the beginning and culminating weeks of the intervention; we use this pre-post comparison of means throughout the analysis in this paper, because the limited length of the intervention renders mid-intervention measurements and predictive analyses non-significant (see [28]). The change response efficiency was significant $M = 1.35, SD = 3.13, N = 26, t(25) = 2.206$, two-tailed, $p < 0.05$ (Figure 4). A Cohen's d statistic was also computed to measure this effect size (Cohen's $d = .57$), which indicates a moderate intervention effect. This supports *H2*, that children's engagement increases over time.

C. Conversation as a Means of Interaction

We were also interested in the level of conversation that the child had with the robot. To demonstrate this, we calculated each participating child's mean length of utterance (MLU). The MLU is a widely utilized proxy for speech production by researchers and practitioners in the speech and language fields. We define an utterance to be a complete response to the robot for the purposes of this calculation. The mean length of utterance for participants across all weeks of the intervention was 28.41 ($SD = 24.57$). There were successive changes in response words to the robot over time. There was a 2.29 mean increase in utterance length from the start to end of the intervention period, in accordance with *H3*, however this trend did not reach statistical significance.

In order to explore differences in the types of utterances employed by the children, we employed an empirical analysis of the spoken language transcripts. We analyzed the content of 137 hand-transcribed verbal responses employed by the participants, categorized these interactions and quantified them (via frequency distribution) to measure changes in types and frequency of interaction over time. The categories we used to identify patterns in the transcripts included: (a) simple responses to prompts, which included one-to-three word responses to robot prompts (e.g., yes, no, huh, um, okay), (b) expansions that included responses with details (e.g., it's healthy, I like it, a magic plate!), and (c) relational responses, which demonstrate evidence that the child was beginning to relate to the robot (e.g., he's hungry? you said you didn't like it). While there was great variability in

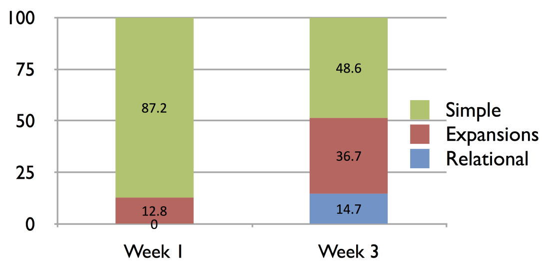

Fig. 5. Response categories over time in child-robot interaction

response types across children, we noted robust changes in types of interactions over time. For example, one child began his interaction with "yeah" and "maybe," and proceeded to recall what the robot had said in a previous session and said "you said you don't like mashed potatoes!", demonstrating both relational speech and, given the tone of the interaction, humor with the robot. We computed the percentage of each category by child and then computed an average percentage per week, per category for the study sample. Figure 5 illustrates these changes between week one and three of the intervention. Thus we find overall some weak support for *H3*, that children use more complex speech with the robot.

D. SAR in the Context of Learning

Similar to our measurement of child-robot interaction, we calculated the mean time that children in the intervention took to make food choices in response to the robot's prompting. The mean response time increased from 13.11 seconds in the first session to 15.36 seconds in the final intervention session, a 2.25 second increase. A paired t-test was performed to ascertain whether the child's food selection time became more efficient (measured by decreases in response times). The mean response difference for food selection ($M = 2.24, SD = 5.26, N = 26$), ($t(25) = 2.174$, two-tailed, $p < 0.05$), Cohen's $d = .45$ providing evidence of the effect size of this change (see Figure 4). This is in direct contradiction to *H4*, however, the child task expectations increased in difficulty each week. Accordingly, we interpreted this result as an indication that as the tasks in the interventions became successively more challenging, the children took additional time to make thoughtful selections.

As an indicator of whether the child participants had learned how to make healthy food selections, we computed the ratio of poor to healthy food selection for each child as a way to normalize choice quality across sessions and facilitate comparison. We found a trend towards improved choices, with a decrease of the mean poor choice ratio from .46 to .39, indicating that the children may have begun to make healthier food selections between the first and last intervention sessions, weakly supporting *H5*, however this change was not statistically significant.

E. Temperament, Interaction, and Learning

As a final comparative measurement, we wished to determine if child temperament and associated behavior influenced child interaction with the robot. Therefore, we used CBQ-S as a comparative metric for our other research results. The mean child self-regulation score on the CBQ-S was 3.08 ($SD = .53$), the mean child positive affect level was 2.95 ($SD = .38$), and the mean negative child affect was scored as 2.40 ($SD = .34$) for our study sample. Specifically, using this measure, we wished to determine whether child temperament predicted child interaction time, child food selection time, or food ratio (described above). Accordingly, each of these variables was selected as dependent variables in multiple regression analysis, with Bonferroni error correction. A step-wise regression was selected using our theoretical perspectives as a means of determining the order in which independent variables would be loaded into the model. Accordingly, socio-demographic variables were loaded into the model first, followed by child temperament variables. These analyses revealed that positive child affect was moderately predictive of food ratio (serving as a proxy for learning), $r^2 = .359, F(3, 14) = 6.15, p < .05$, supporting the first part of *H6*, that children with a positive affect will have greater learning agains. Child-robot interaction and food selection did not contribute to the model and therefore are not predicted by child temperament (thus leaving *H6* unsupported), however child temperament (which contributed to the model) may predict child learning.

V. Discussion

The goal of this study was to examine the feasibility of a SAR-based intervention for teaching children about nutrition. We wished to examine the children's evaluation of, and engagement with, the robot, changes in their verbal interaction with the robot over time, the learning effects of the SAR system, and the effect of child temperament on interaction and outcomes.

We find that children rate the robot highly positively across all measures used, and retain a positive perception of the robot after a three-week intervention. We find that over the course of the intervention, children respond more quickly to the robot's verbal queries, suggesting that they not only maintain engagement with the SAR system over time, but likely become more comfortable with the system, perhaps even building rapport with the robot character. While we do not find changes in their MLU, we do find that they use more complex speech with the robot over time, suggesting that the social presence of the robot encouraged child-robot relationship building. We do not find a relationship between between child temperament and social interaction with the robot (*H6*), but this may mean that children with diverse temperaments can interact equally well with the SAR system.

In terms of the educational goal of the intervention, we find limited evidence that the children in the study learned about nutrition over the intervention (*H5*). We find that positive affect is moderately predictive of healthier food choices (our proxy for learning), which is consistent with the literature on education. Finally, our results indicate that children take longer to make food selections over time, contrary to *H4*, however we have modest evidence that children choose healthier foods over time, suggesting that the increase in time may be indicative of greater thoughtfulness in their responses.

In terms of limitations, while this study examines a relatively large, diverse group of children over many sessions, the sample size and intervention length are still limited relative to typical nutrition interventions. Because this work is designed to examine the changes in the interaction over time, we use a within-subjects design and do not include a control group. In future work with a more primary focus on learning, we would include such a control.

VI. Conclusions

In summary, our results provide a promising basis for future SAR interventions for teaching young children, in particular for the task of nutrition education. We demonstrate that children have an extremely positive reaction to the robot, even after several weeks of periodic interaction with the SAR system. Children were able to sustain engagement with the robot character, and perhaps even build a relationship with the robot, a result that is promising for future, longer, interventions. While we do not find strong learning gains over this intervention, we expect that such longer interventions would show greater changes in knowledge. We are therefore encouraged to extend this work to longer interventions as we move towards a more autonomous SAR system.

ACKNOWLEDGMENT

We would like to thank Michelle Zamora of Viva La Puppet (http://vivalapuppet.com), Brittany Flores, Barry Lumpkin, Jacqueline Kory, Sooyeon Jeong, Florian Pellet, Sigurðður Örn, Akhil Khemani, Amita Parikh, Kelley Le, Peppy Sisay, Dennis Orozco, Prayaas Jain, Fred Wen, Gillian O'Reilly, Clifford Nass, Cynthia Breazeal and the faculty and staff at the school sites.

References

[1] J. F. Jent, L. N. Niec, and S. E. Baker, "Play and Interpersonal Processes," in *Play in Clinical Practice: Evidence-Based Approaches*, S. W. Russ and L. N. Niec, Eds. New York, NY: Guilford Press, 2011.

[2] J. Singer and M. Lythcott, "Fostering school achievement and creativity through sociodramatic play in the classroom," in *Children's Play: The roots of reading*, E. F. Zigler, D. G. Singer, and S. J. Bishop-Joseph, Eds. Washington, DC: Zero to Three Press, 2004, pp. 77–93.

[3] C. L. Ogden, M. D. Carroll, B. K. Kit, and K. M. Flegal, "Prevalence of obesity and trends in body mass index among US children and adolescents, 1999-2010." *J. of the American Medical Association (JAMA)*, vol. 307, no. 5, pp. 483–90, Feb. 2012.

[4] A. S. Singh, C. Mulder, J. W. R. Twisk, W. van Mechelen, and M. J. M. Chinapaw, "Tracking of childhood overweight into adulthood: a systematic review of the literature." *Obesity Reviews : an Official Journal of the Int. Assoc. for the Study of Obesity*, vol. 9, no. 5, pp. 474–88, Sept. 2008.

[5] D. S. Freedman, W. H. Dietz, S. R. Srinivasan, and G. S. Berenson, "The relation of overweight to cardiovascular risk factors among children and adolescents: The Bogalusa heart study," *Pediatrics*, vol. 103, no. 6, pp. 1175–1182, June 1999.

[6] D. Spruijt-Metz, "Etiology, treatment and prevention of obesity in childhood and adolescence: A decade in review," *J. of Research on Adolescence*, vol. 21, no. 1, pp. 129–152, Mar. 2011.

[7] E. B. Tate, D. Spruijt-Metz, G. O'Reilly, M. Jordan- Marsh, M. Gotsis, M. A. Pentz, and G. F. Dunton, "mHealth approaches to child obesity prevention: Successes, unique challenges, and next directions," *Translational Behavioral Medicine*, July 2013.

[8] M. D. Hingle, L. Macias-Navarro, A. Rezaimalek, and S. B. Going, "The use of technology to promote nutrition and physical activity behavior change in youth: A review," *The Research Dietetic Practice Group Digest*, pp. 1–10, 2013.

[9] K. Hieftje, E. J. Edelman, D. R. Camenga, and L. E. Fiellin, "Electronic media-based health interventions promoting behavior change in youth: a systematic review." *JAMA Pediatrics*, vol. 167, no. 6, pp. 574–80, June 2013.

[10] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, "The Physical Presence of a Robot Tutor Increases Cognitive Learning Gains," in *Proc. of the Annual Meeting of the Cognitive Science Society (CogSci)*, no. 1, Sapporo, Japan, 2012, pp. 1882–1887.

[11] J. Movellan, M. Eckhardt, M. Virnes, and A. Rodriguez, "Sociable robot improves toddler vocabulary skills," in *Proc. of the 4th ACM/IEEE Int. Conf. on Human-Robot Interaction*. NY, NY, USA: ACM Press, Mar. 2009, p. 307.

[12] P. K. Kuhl, F.-M. Tsao, and H.-M. Liu, "Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning." *Proc. of the National Academy of Sciences of the USA*, vol. 100, no. 15, pp. 9096–101, July 2003.

[13] F. Tanaka and S. Matsuzoe, "Children Teach a Care-Receiving Robot to Promote Their Learning: Field Experiments in a Classroom for Vocabulary Learning," Jan. 2012.

[14] S. Yun, J. Shin, D. Kim, C. G. Kim, M. Kim, and M.-T. Choi, "Engkey: tele-education robot," in *ICSR'11 Proc. of the Third Int. Conf. on Social Robotics*, ser. Lecture Notes in Computer Science, vol. 7072, Amsterdam, Netherlands, Nov. 2011, pp. 142–152.

[15] I. Leite, R. Henriques, C. Martinho, and A. Paiva, "Sensors in the wild: Exploring electrodermal activity in child-robot interaction," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Mar. 2013, pp. 41–48.

[16] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2008, pp. 3230–3235.

[17] A. Setapan, "Creating robotic characters for long-term interaction," Master of Science, Massachusetts Institute of Technology, 2012.

[18] N. González, L. C. Moll, and C. Amanti, Eds., *Funds of knowledge: Theorizing practices in households, communities, and classrooms*. Mahwah, NJ: Erlbaum, 2005.

[19] J. Ormrod, *Educational Psychology: Developing Learners*, 5th ed. Upper Saddle River, NJ: Merrill, 2006.

[20] C. L. Bottari, C. Dassa, C. M. Rainville, and E. Dutil, "The IADL Profile: Development, content validity, intra- and interrater agreement," *Canadian J. of Occupational Therapy*, vol. 77, no. 2, pp. 90–100, Apr. 2010.

[21] M. K. Rothbart, S. A. Ahadi, K. L. Hershey, and P. Fisher, "Investigations of temperament at three to seven years: the Children's Behavior Questionnaire," *Child Development*, vol. 72, no. 5, pp. 1394–408, 2001.

[22] M. Wilson, *Constructing measures: An Item Response Modeling Approach*. New Jersey: Lawrence Earlbaum, 2013.

[23] M. Lombard, T. B. Ditton, D. Crane, B. Davis, G. Gil-Egui, K. Horvath, and J. Rossman, "Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument," in *Third Int. Workshop on Presence*, Delft, The Netherlands, 2000.

[24] C. Kidd and C. Breazeal, "Effect of a robot on user perceptions," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, vol. 4. IEEE, 2004, pp. 3559–3564.

[25] Y. Jung and K. Lee, "Effects of physical embodiment on social presence of social robots," in *Proc. of the Seventh Int. Conf. on Presence*, Valencia, Spain, 2004.

[26] R. H. Poresky, C. Hendrix, J. E. Hosier, and M. L. Samuelson, "The Companion Animal Bonding Scale: Internal reliability and construct validity," *Psychological Reports*, vol. 60, no. 3, pp. 743–746, June 1987.

[27] A. C. Thomason and K. M. La Paro, "Measuring the Quality of TeacherChild Interactions in Toddler Child Care," *Early Education & Development*, vol. 20, no. 2, pp. 285–304, Apr. 2009.

[28] E. J. Pedhazur and L. P. Schmelkin, *Measurement, design, and analysis: An integrated approach*. Lawrence Erlbaum Associates, 1991.