A Model Child? Behavior Models for Simulated Infant-Robot Interaction

Ameer Helmi, Kristen M. Koenig, and Naomi T. Fitter

Oregon State University (OSU), Corvallis, OR 97331 USA
{helmia,koenigkr,naomi.fitter}@oregonstate.edu

Abstract. Simulated child-robot interaction offers ways to test robot behaviors before real-world trials with vulnerable populations. At the same time, this type of simulation requires realistic models of child behavior. We combined cognitive science research on infant attention with real-world child-robot interaction data to develop two behavior tree-based models of child behavior (i.e., robot-interested and robot-uninterested). We evaluated these models through a video-based study ($N = 60$). Participants rated the proposed models as more familiar, humanlike, and natural than a control (random behavior) model. This work can support related work on child-robot interaction, as well as broader efforts on using technology to support infant development.

1 Introduction

Worldwide, approximately half of children between the ages of two and six achieve the recommended amounts of physical activity [24]. Early interventions that encourage physical activity and motor exploration are vital to promoting the interrelated development of physical, cognitive, and social skills [6,15]. Assistive child-robot interventions are gaining attention due to the potential of robots to dovetail with early intervention services. For example, a NAO and Dash robot were combined with a body-weight support harness to encourage movement by a child with Down's Syndrome [11]. Other teams used robots to encourage leg motion practice in infants [3,5]. We designed a custom assistive robot to encourage motor exploration in children [25]. The robot, shown in Fig. 1, comprises a TurtleBot2 base and a custom-designed reward module capable of supplying developmentally appropriate stimuli for young children.

In past work, we conducted an exploratory study with children in free ambulatory play and discovered that our assistive robot effectively encouraged children to stand up and move [25,26]. We plan to build on these promising findings by studying a range of robot planning strategies for encouraging infant motion; however, young children are a vulnerable population. Further, work in the space of infant-robot mobility interventions often involves brief studies (i.e., as short at eight minutes per session) with small numbers of participants due to challenges keeping very young children on task. Thus, situated child-robot interaction data is extremely limited, and we require methods for simulating sufficiently realistic child-robot interactions to ensure that assistive robots are as safe, reliable, and viable as possible before real-world deployments [20]. In the limited work

Fig. 1: Child behavior from past studies. *Left:* Custom assistive robot with light, sound, and bubble stimulus hardware. *Center:* Robot-interested child interacting with our assistive robot. *Right:* Robot-uninterested child playing with toys.

addressing models of child-robot interaction, one framework that has been used is ACT-R, which is similar to a programming language for designing models of simulated human cognition [2]. In past work, ACT-R modeled a single child's actions during hide-and-seek with a robot [23]. Other efforts used Hidden Markov Models to model infant free-play behavior, focusing on toy selection [14]. Our method builds on this past work by combining psychology research on child attention with annotated child-robot interaction video data to construct more realistic models of child behavior. The key research goal of this work was to *design and evaluate a beginning set of child behavior models that can help us vet robot interaction strategies for early mobility interventions.*

In this paper, we first describe the two proposed infant behavior models (i.e., robot-interested and robot-uninterested) and how we constructed the behavior trees for each model (Section 2). In Section 3, we discuss our online video-based evaluation of the models, and the results of the evaluation appear in Section 3.2. Section 4 discusses the implications of our models and offers directions for future research. The main contribution of this paper is the design and initial validation of two data-driven behavior tree models for infant behavior in simulated child-robot interaction.

2 Infant Behavior Models

Cognitive science research informed our methodology for creating infant behavior models. We annotated video data from a prior exploratory playgroup and constructed behavior tree models for two simulated infants.

2.1 Cognitive Science Foundations

Cognitive science methods for video coding and past findings related to child attention informed our efforts to model infant behavior. Lansink et al. performed seminal work on annotating child attention during interactions with objects [12]. Video annotators in the study used behavioral labels of casual and focused attention to note the child's attention. Casual attention was marked by general inattentiveness and a high frequency of looking away from objects [13,18]. Focused attention was marked by a decrease in heart rate, longer glances, and a lower likelihood of looking away from the object of interest [12,19]. In our own video coding, as described further in Section 2.2, we adopted these types of attention as the core of our annotation strategy.

Additionally, past cognitive science work shows that infants cycle between casual attention and focused attention and spend measurable characteristic durations attuned to different objects of interest (e.g., toys) [12]. When a child is in focused attention, Lansink et al. found that the child will stay focused on an object 96.5% of the time and look away 3.5% of the time [12]. We formed our models, as further explained in Section 2.3, partly based on this information from past literature and partly from our own annotations.

2.2 Playgroup Data Annotation

Our team previously recorded overhead video data from a study during which children interacted with our assistive robot and developmentally appropriate toys in an open play space, as further reported in [25]. The assistive robot is composed of a TurtleBot2 base (a common mobile robotic platform capable of non-holonomic base motion in 2D) and a custom reward module for delivering stimuli (e.g., LED light patterns, sounds, bubbles) to child users of the system. We selected the 30-minute video of the children's initial play behaviors with our robot for analysis since this segment captured ad-hoc interaction; the robot is intended for early interventions that may be one-time sessions.

Specifically, we annotated the behaviors of the two youngest children in the study (1.5 and 2 years of age), as these participants best fit the age range for the types of early interventions that our robot was designed for (i.e., below 3 years of age). We observed that the younger child displayed a limited amount of time with the robot, and accordingly, formed the *robot-uninterested* model based on this infant. The older child spent a moderate amount of time playing with the robot and thus provided the foundation for the *robot-interested* model. Example interactions by these two children appear in Fig. 1.

From the video of these two children, we annotated periods of casual and focused attention, as well as the duration of time focused on the robot and toys. We excluded coding times when either child was out of the camera field of view. Based on the cognitive science groundings from the previous subsection, we used the following codebook to annotate the video:

- *Casual attention:* shifting gaze continually or failing to focus on clear target.
- *Focused attention on a toy:* performing primarily long glances at toy(s) or playing directly with toy(s).
- *Focused attention on the robot:* performing primarily long glances at robot or playing directly with robot.
- *Interaction duration:* length of time that one of the above states lasted.

The results of our video annotation for each child appear in Table 1. The robot-interested child spent a larger proportion of time in focused attention with both the robot and toys in the play space. Additionally, the robot-interested child spent a higher mean time playing with the robot and toys. We applied these results to the designed behavior tree models of infant behavior, as further explained in the following subsection.

Table 1: Video annotation results including percentages of infant casual and focused attention, as well as mean and standard deviation time spent with items of interest when in focused attention.

Behavior Code	Robot-Interested	Robot-Uninterested
Casual Attention	51.7%	81.1%
Focused Attention - Toy	17.5%	11.4%
Focused Attention - Robot	30.8%	6.8%
Interaction Duration - Toy (s)	19.1 ± 15.9	12.5 ± 10.4
Interaction Duration - Robot (s)	35.1 ± 34.0	15.0 ± 10.1

2.3 Behavior Trees

Behavior trees have been used to control autonomous agents in video games [7] and for supporting robot decision-making [1,16], but have yet to be applied to modeling infant behavior. We identified behavior trees to be a viable model option since they are easy to interpret and allow for flexibility in incorporating future data [4]. In this subsection, we outline background on behavior trees and describe our proposed behavior tree models of child behavior. When discussing the basic functions of behavior trees, we use the standard terminology of parent and child node, not to be confused with a human parent and human child.

A behavior tree operates as a top-down left-to-right hierarchical tree that evaluates true or false *condition* nodes in the tree and determines an *action* node to activate. Internal nodes, such as *fallback* and *sequence* nodes, are used to control the flow of the tree [4]. Branches of the behavior tree will return one of three statuses to the root node: *success* if a node completes, *failure* if a condition or action fails, or *running* if an action is in progress. *Fallback* nodes, represented with a ?, are used when only one branch underneath the fallback node should be active. If one child node of the fallback returns *success* or *running*, the fallback node will return *success* to its parent node. *Sequence* nodes, represented with an arrow ($\rightarrow$), are used when all child nodes under the sequence node should return *success* or *running*. A sequence node returns *success* to its parent node only if all child nodes return *success*; otherwise, it will return *running* or *failure*. Statuses are returned by child nodes to parent nodes until reaching the root (topmost) node, which indicates the current status of the entire behavior tree. Other node types, such as parallel nodes, exist in behavior trees broadly but are not used in our implementation.

We developed two behavior trees (robot-interested and robot-uninterested models) for determining a simulated child's actions while interacting with toys and our robot. All branches of the behavior trees are available for viewing on our public repository [9]. Our behavior trees operates in three distinct layers: the *visual field* layer, the *attention type* layer, and the *action* layer. Each layer has conditions which determine which branch of the next layer to proceed to. We outline each layer and the associated condition or action options below:

- Visual field layer
  - See robot + toy(s), See toy(s), See robot, or See neither robot nor toy(s)
- Attention type layer

    - Focused attention on toy, Focused attention on robot, or Casual attention
- Action layer
    - Play with toy, Play with robot, Stand still, or Move random direction

Each branch of the tree begins by evaluating what objects are visible to the child, e.g., if the child sees only the robot, then the probability of focused attention on the toy is removed. The conditions in the *attention type* layer are probabilities based on the percentages from Table 1. The conditions in the *action* layer are probabilities based on percentages described previously in [12], i.e., a child will stay focused on an object 96.5% of the time and look away 3.5% of the time. Each action (e.g., Play - Toy) in the *action* layer is performed for a specific length of time (based on the mean lengths of time shown in Table 1) and returns a status of *running* during the action. A child that is not playing with either a toy or the robot will move in a random direction for two seconds or stay still for two seconds. Once an action is completed, the node returns *success* and the tree is re-evaluated for a new action.

3 Model Evaluation

We conducted an online within-subjects study to assess how the child behavior models were perceived relative to one another, as well as relative to a control (random behavior) model. All study procedures were approved by Oregon State University under protocol #IRB-2019-0172.

3.1 Methods

The evaluation presented participants with three different videos depicting interactions between a simulated child and surroundings including multiple toys and a robot. Each video represented one condition in our evaluation:

- *Robot-Interested:* child behavior followed robot-interested model.
- *Robot-Uninterested:* child behavior followed robot-uninterested model.
- *Random:* child has an equal probability of each type of attention described previously, and time spent in each action is randomized.

Video Stimuli: We developed a custom simulation using Python3 and ROS Noetic. The simulated child was placed in a virtual play space with three randomly placed static toys and the assistive robot. A blue line originating from the child indicated gaze direction to participants. The child's field of view was 60 degrees from each side of the blue line and determined which objects were visible at each time step; this field of view was not conveyed to participants. The simulated child's image would flip between facing left or right to match the gaze direction, as shown in Fig. 2. The simulated child acted according to one of the conditions and could move in the environment and interact with toys or the robot. Each condition, including random, was programmed in Python3 as a behavior tree which evaluated and chose actions during each time step. ROS Noetic was used to control the timings of behaviors during the simulation and to allow for future integration with behavior tree packages. The robot behaved according to a uniform behavior tree across conditions, the purpose of which is

Fig. 2: *Left:* Stimulus frame depicting the location of the mock infant, toys, and robot. The blue line originating from the child indicates gaze direction. *Right:* Stimulus frame showing the robot bubbles action. The child image orientation has flipped to match the gaze direction.

to stay near the child to encourage play. The generated video stimuli were 45 seconds on average. Fig. 2 shows two frames from one stimulus video.

Procedure: Participants were recruited via a university student pool. Respondents completed the study online via a Qualtrics survey that began with an informed consent page. Next, participants completed demographic questions and read introductory information that explained the simulated interaction scenario. In each of the following survey blocks, respondents watched one of the three randomly ordered video stimuli and responded to questions about the simulated child after the video concluded. Finally, participants completed an attention-check question and described the factors influencing their survey responses via a free-responses question requiring a minimum of 200 characters. Participants received course credit for completing the study.

Measures: The survey included questions about the simulated infant's behavior and apparent interests, as well as basic demographic questions. Five survey questions were administered on a seven-point Likert scale, as described below:

- *Familiarity* of child behavior, "Very Strange" (1) to "Very Familiar" (7)
- *Humanlikeness*, "Very Non-Humanlike" (1) to "Very Humanlike" (7)
- *Naturalness* of behavior, "Very Artificial" (1) to "Very Natural" (7)
- Whether the child was *interested in the toys*, "Strongly Disagree" (1) to "Strongly Agree" (7)
- Whether the child was *interested in the robot*, "Strongly Disagree" (1) to "Strongly Agree" (7)

The first two questions above came from inventories on the uncanny valley [10], and the *naturalness* question is our own exploratory addition. The final scales sought to capture perceived differences in the simulated infants' interests. Further demographic questions gathered information on participant age, identity, profession, experience with children, and experience with robots.

Hypotheses: Our two main hypotheses were as follows:

- **H1**: Participants will rate the robot-interested and robot-uninterested conditions as higher in familiarity, naturalness, and humanlikeness than the

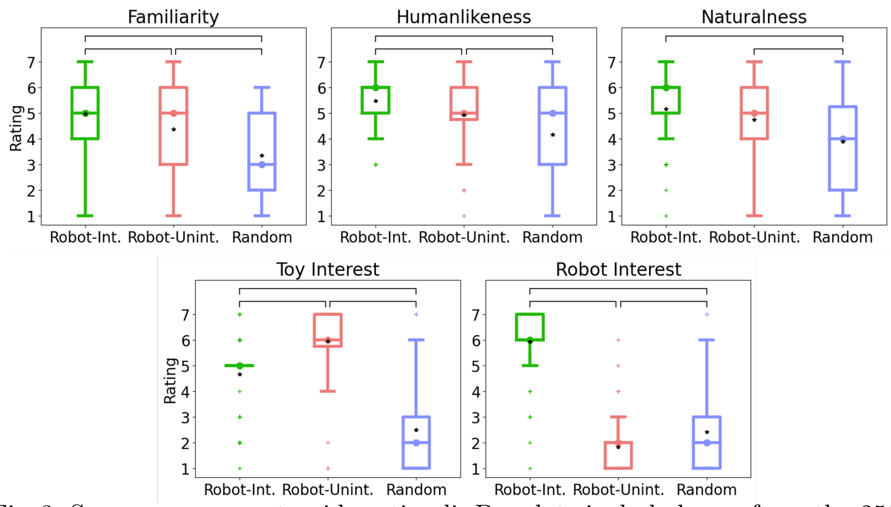

Fig. 3: Survey responses to video stimuli. Boxplots include boxes from the 25th to the 75th percentiles, center lines with a circle marker for medians, asterisks for means, whiskers up to 1.5 times the inter-quartile range, and "+" marks for outliers. Brackets above the boxplots indicate significant pairwise differences.

random condition. In other words, we expect the models founded on real observations of human behavior to seem more lifelike.

- **H2**: Participants will rate the robot-interested condition as most interested in the robot and the robot-uninterested condition as most interested in toys.

These hypotheses arose from the foundation of the model design in cognitive science and annotation of two real children's interactions with our robot.

Participants: The study was completed by $N = 60$ students from Oregon State University between the ages of 18 and 41 ($M = 19.8$ and $SD = 3.3$), including 17 cisgender men, 42 cisgender women, and one non-binary individual. One participant was a parent/guardian and 25 participants indicated experience interacting with children in our early intervention age range of interested (6-36 months). 53 participants indicated little or no experience with robots.

Analysis: We analyzed stimulus video ratings using repeated-measures ANOVA (rANOVA) tests with $\alpha = 0.05$. Tukey's HSD Test for multiple comparisons were conducted for significant main effects. We report effect sizes using η^2, where $\eta^2 = 0.010$ is considered a small effect, $\eta^2 = 0.040$ a medium effect, and $\eta^2 = 0.090$ a large effect [8]. Statistical analyses were conducted using jamovi [17,21,22].

3.2 Results

We compared ratings across the study conditions. Response distributions and rANOVA results are shown in Fig. 3.

Ratings of Familiarity, Humanlikeness, and Naturalness: rANOVA results revealed that there were significant differences in *familiarity* ($p<0.001$, $F(2, 118) = 25.4$, $\eta^2 = 0.171$), *humanlikeness* ($p<0.001$, $F(2, 118) = 18.6$, $\eta^2 = 0.125$), and *naturalness* (p<0.001, F(2, 118) = 16.4, $\eta^2 = 0.112$) ratings. The robot-interested and robot-uninterested models were significantly more familiar, humanlike, and natural than the random model. Participants also rated the robot-interested model as significantly more familiar and humanlike than the robot-uninterested model. We also conducted an exploratory factor analysis of our three main measures. The resulting Cronbach's alpha value showed that these scales may hold promise as a realism construct ($\alpha = 0.92$).

Ratings of Child Interest in Toys and Robot: Significant differences appeared in the *interest in toys* (p<0.001, F(2, 118) = 89.3, $\eta^2 = 0.477$), and *interest in robot* (p<0.001, F(2, 118) = 169, $\eta^2 = 0.646$) ratings. The robot-uninterested model appeared more interested in toys than both other models. The robot-interested model also was rated as more interested in toys than the random model. Participants rated the robot-interested model as significantly more interested in the robot than any other model. The random model also appeared to be more interested in the robot than the robot-uninterested model.

4 Discussion

This paper proposed and evaluated two models of infant behavior based on insights from cognitive science and analysis of real-world interaction data. The model evaluation results support **H1**; participants rated the robot-interested and robot-uninterested conditions higher in familiarity, humanlikeness, and naturalness when compared to the random condition. Likewise, we found support for **H2**. Participants rated the robot-interested model as more interested in the robot and the robot-uninterested model as more interested in the toys, and both findings had very large effect sizes.

Free-text responses further elucidated respondents' model expectations and perceptions. Most participants noted that they focused on what objects drew child interest and how long object interactions lasted. One participant wrote "if the child was playing with toys and the robot, I saw that as less artificial." Another respondent mentioned that "in [one] video, the child seemed scared of the robot and appeared to seek the bear for comfort, [which] seemed like an appropriate response." A comment on the random condition labeled it as "strange" and mentioned that the child "wandered all over the place with maybe a couple of glances towards objects." A further participant noted that "[their] experience [led them] to believe that a child that age should be mesmerized by the toys and especially moving ones such as the robot." Consistently with the cognitive science literature, written feedback supported the idea that it is natural for children to fixate on and engage with items in the environment.

The *strengths* of this work include the extension of related work to effectively simulate infant behaviors during interactions with toys and a robot. The models fuse research in infant attention and video-annotated data of real child-robot interactions. The model evaluation supported that these models function as intended. The current robot-interested and robot-uninterested models, in addition to future models to be created using similar methods, will allow us (and others with similar interests) to vet robot planning algorithms before deploying these strategies in resource-intensive child-robot interactions in the real world.

Limitations of this work include the current scope; namely, the models represent ad hoc interactions with a robot, and the model evaluation centered on just one set of video stimuli. We could follow the same process as used in this paper to annotate further types of child behavior, such as interactions with parents or other children, and build and test similar models more broadly. Other limitations arose from the demographics of respondents. Most participants did not have children of their own, although nearly half of the group had experience with children below three years of age. We could recruit a broader set of participants, including more parents, in future model evaluations. We also plan to use our model to compare real vs. simulated child behavior in future study trials.

In conclusion, this paper presents infant behavior models that use methods from cognitive science, are data-driven, and employ easy-to-read and flexible behavior trees. We demonstrated that our models can produce more realistic-seeming infant behavior than a random model. This work can inform assistive robot designers and others who are interested in modeling infant behavior.

Acknowledgments

We thank Emily Scheide and Ryan Quick for helping to create the study stimuli.

References

1. Abiyev, R.H., Günsel, I., Akkaya, N., Aytac, E., Çağman, A., Abizada, S.: Robot soccer control using behaviour trees and fuzzy logic. Procedia Computer Science **102**, 477–484 (2016)
2. Anderson, J.R., Matessa, M., Lebiere, C.: ACT-R: A theory of higher level cognition and its relation to visual attention. Human–Computer Interaction **12**(4), 439–462 (1997)
3. Boyd, J., Harris, J., Smith, M., García-Vergara, S., Chen, Y.P., Howard, A., et al.: An infant smart-mobile system to encourage kicking movements in infants at-risk of cerebral palsy. In: Proc. of the IEEE Workshop Advanced Robotics and Its Social Impacts (ARSO). pp. 1–5 (2017)
4. Colledanchise, M., Ögren, P.: Behavior trees in robotics and AI: An introduction. CRC Press (2018)
5. Fitter, N.T., Funke, R., Pulido, J.C., Eisenman, L.E., Deng, W., Rosales, M.R., Bradley, N.S., Sargent, B., Smith, B.A., Mataric, M.J.: Socially assistive infant-robot interaction: Using robots to encourage infant leg-motion training. IEEE Robotics & Automation Magazine **26**(2), 12–23 (2019)
6. Fox, S.E., Levitt, P., Nelson III, C.A.: How the timing and quality of early experiences influence the development of brain architecture. Child Development **81**(1), 28–40 (2010)
7. Fu, Y., Qin, L., Yin, Q.: A reinforcement learning behavior tree framework for game AI. In: Proc. of the Int'l Conference on Economics, Social Science, Arts, Education and Management Engineering (2016)
8. Funder, D.C., Ozer, D.J.: Evaluating effect size in psychological research: Sense and nonsense. Advances in Methods and Practices in Psychological Science **2**(2), 156–168 (2019)

9. Helmi, A., Koenig, K.M., Fitter, N.T.: Infant behavior trees repository (July 2022), https://github.com/shareresearchteam/infant_simulation_behavior_trees
10. Kamide, H., Kawabe, K., Shigemi, S., Arai, T.: Development of a psychological scale for general impressions of humanoid. Advanced Robotics **27**(1), 3–17 (2013)
11. Kokkoni, E., Mavroudi, E., Zehfroosh, A., Galloway, J.C., Vidal, R., Heinz, J., Tanner, H.G.: GEARing smart environments for pediatric motor rehabilitation. Journal of Neuroengineering and Rehabilitation **17**(1), 1–15 (2020)
12. Lansink, J.M., Mintz, S., Richards, J.E.: The distribution of infant attention during object examination. Developmental Science **3**(2), 163–170 (2000)
13. Lansink, J.M., Richards, J.E.: Heart rate and behavioral measures of attention in six-, nine-, and twelve-month-old infants during object exploration. Child Development **68**(4), 610–620 (1997)
14. Le, H., Hoch, J.E., Ossmy, O., Adolph, K.E., Fern, X., Fern, A.: Modeling infant free play using hidden markov models. In: Proc. of the IEEE Int'l Conference on Development and Learning (ICDL). pp. 1–6 (2021)
15. Logan, S.W., Schreiber, M., Lobo, M., Pritchard, B., George, L., Galloway, J.C.: Real-world performance: Physical activity, play, and object-related behaviors of toddlers with and without disabilities. Pediatric Physical Therapy **27**(4), 433–441 (2015)
16. Paxton, C., Ratliff, N., Eppner, C., Fox, D.: Representing robot task plans as robust logical-dynamical systems. In: Proc. of the IEEE/RSJ Int'l Conference on Intelligent Robots and Systems (IROS). pp. 5588–5595 (2019)
17. R Core Team: R: A language and environment for statistical computing (Version 4.0) [Computer software] (2020), https://cran.r-project.org/
18. Ruff, H.A.: Components of attention during infants' manipulative exploration. Child Development pp. 105–114 (1986)
19. Ruff, H.A., Lawson, K.R.: Development of sustained, focused attention in young children during free play. Developmental Psychology **26**(1), 85 (1990)
20. Salem, M., Lakatos, G., Amirabdollahian, F., Dautenhahn, K.: Towards safe and trustworthy social robots: Ethical challenges and practical issues. In: Proc. of the Int'l Conference on Social Robotics (ICSR). pp. 584–593. Springer (2015)
21. Singmann, H.: afex: Analysis of factorial experiments [R package] (2018), https://cran.r-project.org/package=afex
22. The jamovi project: jamovi (Version 1.6) [Computer software] (2020), https://www.jamovi.org
23. Trafton, J.G., Schultz, A.C., Perznowski, D., Bugajska, M.D., Adams, W., Cassimatis, N.L., Brock, D.P.: Children and robots learning to play hide and seek. In: Proc. of the ACM SIGCHI/SIGART Conf. on Human-Robot Interaction. pp. 242–249 (2006)
24. Tucker, P.: The physical activity levels of preschool-aged children: A systematic review. Early Childhood Research Quarterly **23**(4), 547–558 (2008)
25. Vinoo, A., Case, L., Zott, G.R., Vora, J.R., Helmi, A., Logan, S.W., Fitter, N.T.: Design of an assistive robot for infant mobility interventions. In: Proc. of the IEEE Int'l Conference Robot and Human Interactive Communication (RO-MAN). pp. 604–611 (2021)
26. Vora, J.R., Helmi, A., Zhan, C., Oliviares, E., Vu, T., Wilkey, M., Noregaard, S., Fitter, N.T., Logan, S.W.: Influence of a socially assistive robot on physical activity, social play behavior, and toy-use behaviors of children in a free play environment: A within-subjects study. Frontiers in Robotics and AI (2021)