5th International Conference on Development and Learning
and on Epigenetic Robotics (ICDL-EpiRob)
Aug 13-16, 2015. Providence, RI USA

Designing a Socially Assistive Robot for Personalized Number Concepts Learning in Preschool Children

Caitlyn Clabaugh*, Gisele Ragusa†, Fei Sha*, and Maja Matarić*
*Computer Science Department
†Division of Engineering Education
Viterbi School of Engineering, University of Southern California, Los Angeles, CA 30332–0250
{clabaugh,feisha,ragusa,mataric}@usc.edu

***Abstract*—Designing technological systems for personalized education is an iterative and interdisciplinary process that demands a deep understanding of the application domain, the limitations of current methods and technologies, and the computational methods and complexities behind user modeling and adaptation. We present our design process and the Socially Assistive Robot (SAR) tutoring system to support the efforts of educators in teaching number concepts to preschool children. We focus on the computational considerations of designing a SAR system for young children that may later be personalized along multiple dimensions. We conducted an initial data collection to validate that the system is at the proper challenge level for our target population, and discovered promising patterns in participants' learning styles, nonverbal behavior, and performance. We discuss our plans to leverage the data collected to learn and validate a computational, multidimensional model of number concepts learning.**

I. Introduction

Socially Assistive Robotics (SAR) is an emerging genre in sociotechnical systems for education. Broadly, SAR applies embodied computation to supplement the efforts of educators, parents, and clinicians in mitigating critical societal problems that require socially mediated, personalized, and long-term support [1][2]. Specifically, SAR tutors aim to supplement the efforts of educators through engaging children in personalized educational activities. SAR presents a unique class of educational tool as it is scalable, lending itself to broad replication as technology advances and become more affordable, as well as socially mediated and embodied, supporting best practices and socio-constructivist learning theories.

Personalization of instruction is essential to enable a child to reach his/her full learning potential. From a computational standpoint, personalization may be scoped to depend on one or many dimensions of a child and his/her interaction with a system. For instance, a simple personalized model may depend solely on whether a child correctly or incorrectly answered a given question, while a multidimensional model may depend on a large variety of factors including the child's learning style, special needs, multiple intelligences, grit, etc., in addition to his/her performance.

It is an important and significant challenge to enable a robot tutor to personalize its model of an individual child. Equally challenging is the design of a system and data collection through which such a model may be learned. The scale of this challenge is strongly dependent on the number and nature of the dimensions or attributes over which the robot is personalizing. While it may be intuitive that a human educator should view a child's ability from as holistic a perspective as possible, computationally, each attribute included in a robot's understanding of a child introduces more perceptual noise and greater computational complexity approaching intractability. Thus, as in most computational modeling problems, *system design*, attribute or *feature selection*, and *statistical method* are intertwined and of equal importance to the success of the educational SAR system. In this paper, we focus specifically on system design and preliminary feature selection, looking towards future statistical modeling.

This paper outlines a set of iterative design processes and considerations necessary to construct a SAR system and data collection that can be used for multidimentional personalization in future work. We present our methodological processes in the context and scope of a novel autonomous SAR system for preschool number concepts education. Our contributions are as follows:

- SAR approach to support number concepts learning in preschoolers through play;
- Design process and considerations behind projecting general classroom number concepts standards into a SAR context;
- Multimodal data collection to capture child performance and nonverbal behavior;
- Preschooler learning style inventory questionnaire;
- Preliminary analysis of childrens performance, learning style, and nonverbal behavior.

We first summarize key components of the extensive foundational research upon which our work rests; then we describe the SAR system and our design and data collection processes, followed by some preliminary analysis of the data collected and a brief discussion of implications for future work.

II. Background

Science, technology, engineering, and mathematics (STEM) is currently a major thrust in education because of market demands [3]. Academic competence in mathematics is both critical [4] and achievable [5] in early childhood education. Thus, number concepts, the understanding of number basics [6], is a vital developmental goal in preschool education. However, most ITS research for personalized STEM education

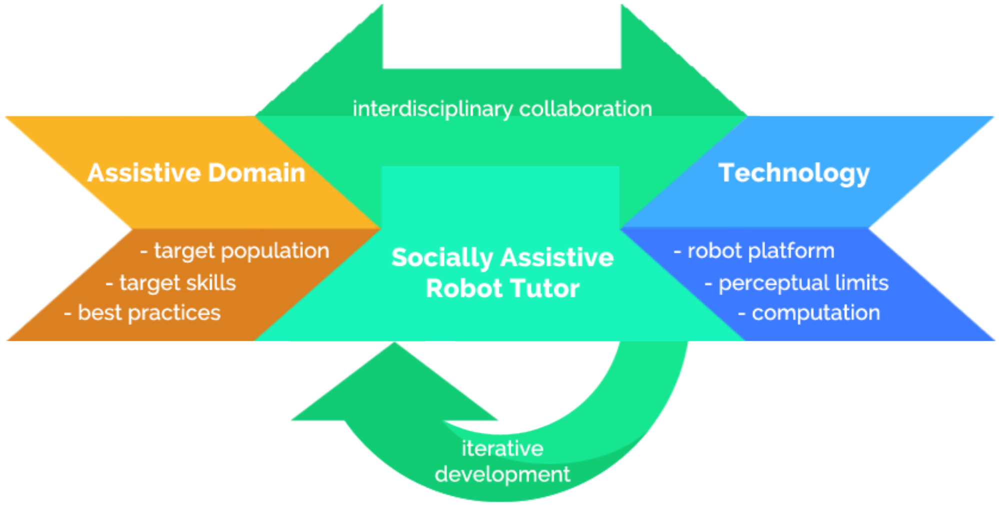


Fig. 1: A visualization of the bidirectional, interdisciplinary, and iterative SAR design process.

has been studied and designed for undergraduate education [7][8][9][10]. Thus, we are faced with the unmitigated challenges of projecting educational material onto a new sociotechnical system with developmentally appropriate practices.

Given that we are interested in designing intelligent tutoring systems for preschool children, SAR provides a unique set of advantages as a profoundly interactive and co-present learning paradigm [11]. The embodiment of SAR has shown to increase cognitive learning gains [12] and compliance [13]. Additionally, the intrinsic agency of a SAR lends itself to learning through play [14], a method important to socio-constructivist theories of early-childhood learning [15][16]. Evidence for sustained engagement, at least among some children, has been demonstrated with a SAR system [17]. For these reasons, we are driven to develop personalized STEM tutoring in a SAR context.

As outlined in the introduction, we are also motivated to investigate multimodal and affective features of the child's interaction. Most ITS personalization considers only the correctness of students' answers, leaving affective and multimodal data entirely untouched. This is also true in the case of SAR educational intervention research [18][19]. In a paradigm that is so deeply grounded in social interaction and inspired by socially mediated education, it is important to expand the evaluation of student ability to more than one modality.

In addition to real-time, multimodal features of performance and interaction, we are interested in exploring and leveraging learners' individual differences, specifically *learning styles*. Learning styles aim to classify the manner by which individuals approach learning with possible impacts on their performance (see [20] for a review). There exists little to no research on learning styles in preschool children. Therefore, one of the main thrusts of our research was to define a preschool learning styles model and corresponding inventory questionnaire for parents and teachers. We based our model off the existing, well-researched work of Felder and Silverman for undergraduate engineering students [21]. Felder's model defines five dimensions [21] from which we have selected the three most perceptible dimensions in preschool children: visual, verbal, and active. We describe our model and questionnaire in greater detail in the following Section III-A.

We present a hybrid approach through which the learning styles of the child participants and the exercises are defined by literature-based features (i.e., parent and teacher questionnaires and learning expert characterization of exercises), and multimodal features of interaction can be extracted and analyzed computationally. Learning styles are as innate or impermanent to the learner as to the learning situation. In our approach, we define a closed set of educational exercises, bearing in mind that each exercise may have a static, inherent, and prefered learning style and each child participant's learning style is dynamic and influential.

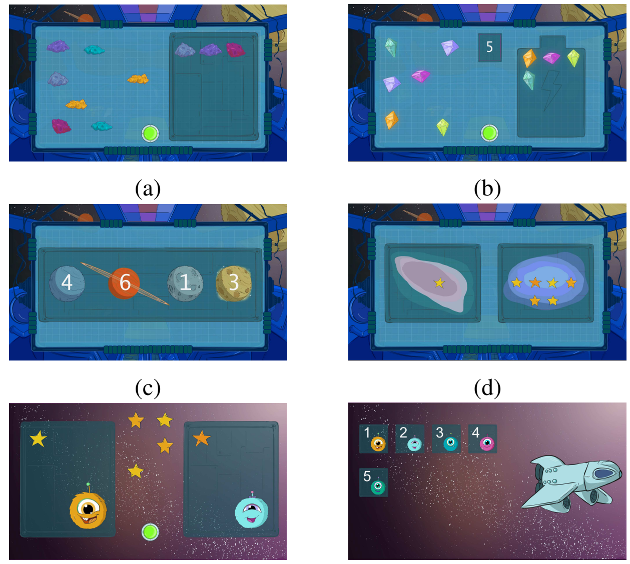
III. Approach

The development of personalizable SAR is a bidirectional interdisciplinary process through which the demands of an assistive domain and the limitations of current technologies must come into equilibrium (see Figure 1). In the following section, we describe in detail our iterative development process for a personalizable SAR approach to learning in the domain of preschool number concepts.

A. System Design & Implementation

The design and implementation of the SAR system, presented in this section, considers expert knowledge across the three principal disciplines: education, machine learning, and robotics. We iteratively and carefully balanced the educational goals with the computational and engineering challenges necessary for future personalized statistical modeling. Initially interested in the domain of STEM education for young children, we found that number concepts minimized the perceptual requirements of system such that it is *fully autonomous* and targets a critical skill in early childhood STEM education.

Fig. 2: The SAR system setup in preschool classroom. The system consists of a Aldebaran NAO robot and a touch-enabled tablet. All system content is openly available for research purposes



(a) (b) (c) (d) (e) (f)

Fig. 3: Screen captures of activity types.

Based on the 2013 Transitional Kindergarten Math Pacing Guide [22], we designed a set of seven unique exercise types that require only touch or click input from the child. Each exercise type can be configured for numbers one through nine. Through a space-themed background story, the robot acts as a knowledgeable peer in a playful interaction centered around the number concepts exercises. The robot is a "robot space-explorer" named "Vega" and needs the child's help to accomplish a set of tasks so that it can return home, encouraging the child to learn through play [14]. The child's performance does not influence the overall outcome of the game ("Vega" is always able to return home) and is designed such that the child continues to the next activity even if he/she cannot answer correctly. The seven types of exercises are as follows:

1) "Put $[1, \ldots, 9]$ moon rocks into the box.", see Figure 3a.
2) "Put ['most', 'all', 'just one'] moon rocks into the box., see Figure 3a.
3) "What number do you see? Put that many energy crystals into the battery pack...", see Figure 3b.
4) "Choose the planet with the number $[1, \ldots, 9]$.", see Figure 3c.
5) "Choose the galaxy that has ['more', 'less'] stars.", see Figure 3d.
6) "Feed Uki and Yana the same amount of stardust.", see Figure 3e.
7) "Move the space pets into the spaceship in order, from 1 to $[2, \ldots, 9]$. Count them out loud with me!", see Figure 3f.

These exercises were iteratively developed through a series of focus groups with teachers and beta tests with children from the target population. Through this process, we incorporated developmentally appropriate practices, minimal perceptual complexity, and expert feedback from target population into our design.

Additionally, our design takes into account the need for computational representations of the exercises for future statistical modeling. Naturally, each exercise has an inherent learning style. Our computational characterization of the exercises considers learning style such that the learning style may correlate with performance on exercise type. We characterize each of the seven exercise types for number $n \in [9]$ as some combination of the following features:

- **Magnitude** – requires the magnitude of n be known
- **Order** – requires the order of numbers 1 to n be known
- **Name** – requires the name of n be known
- **Symbol** – requires the numeric symbol for n be known

To control for optimality of engaging peer-like presentation, we had a graphic artist create the game artwork, and used a pre-recorded enthusiastic human child-like voice for the robot's speech, recorded in both English and Spanish. The robot used deictic gestures to guide the child's motions on the tablet interface. We chose to implement each of these tasks manually as each constitutes its own large research area outside of the scope of our work, and thus, this work does not investigate such specific components of human-robot interaction. In this regard, we again relied heavily on interdisciplinary collaboration in developing the final SAR system.

System portability was a significant design consideration. While we concluded that on-board robot sensing does not provide the breadth of multimodal data we needed (i.e., the cameras onboard the Aldebaran NAO only provide one view of the child, eliminating the possibility for postural and tablet-interaction data), our system deliberately contains a set of a sensors that are easy to port and set up. Thus, it is feasible to

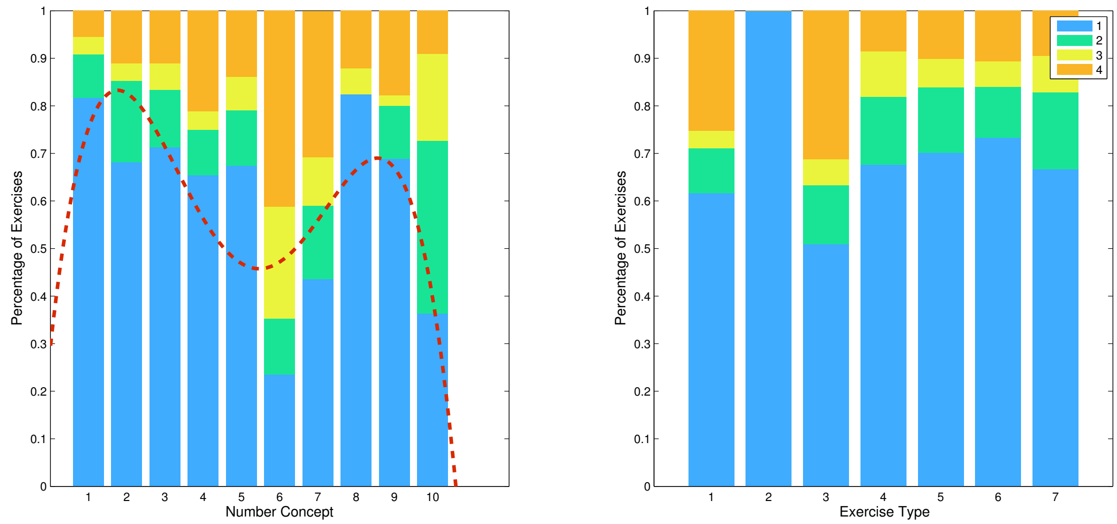


Fig. 4: The distribution of the number of attempts made by the child participants over the number concepts and exercise types. 1, 2, and 3 signify the child's first, second, and third attempts, respectively, and 4 signifies the robot moved onto the next exercise. Children's performance is bimodal wrt exercises' number concepts, and near uniform wrt to exercises' types.

use our system in a general preschool classroom. In addition to their portability, the set of sensors lends itself to online feature extraction for future multidimensional personalization.

We collected a variety of multimodal data from each participant, including: *a*) front-facing audio/visual data from a high definition camera for head position and facial landmarks, *b*) profile visual data from webcam for postural information, *c*) touch-based tablet interactions, and *d*) overhead visual data from a webcam for any hand movements that did not activate the tablet touch tracking. This multimodal dataset lends itself to a more comprehensive analysis of the child's interaction while remaining viable as a data collection setup in real-world classroom environments.

In addition to an individual child's performance on exercises and nonverbal interaction with the SAR system, we observed semi-static features of a child's character, specifically learning style. We constructed a learning styles inventory questionnaire for teachers and parents to obtain some qualitative measure of an individual child's learning style defined by the child's preference of visual, verbal, and active learning. The exercises and learning style questionnaire were designed in tandem to maximize the probability of correlations between child's learning style and performance.

Given the above design considerations, we believe our SAR approach gives equal importance to the standards, advancement, and demands of the domain and the technology. Our approach achieves full autonomy, multimodality, robust functionality, and computational rigor, while preserving portability, practicality, and developmental relevance to a general education classroom. We encourage the same iterative, interdisciplinary development process for other SAR research targeting personalized education.

B. Data Collection

The majority of our efforts were spent on the iterative development of a robust, portable SAR system that is developmentally appropriate for preschool number concepts learning with a rigorous computational representation. As portrayed in the previous sections, our design process was not conducted in isolation. We developed our system in collaboration with experts in education and through a series of informal focus groups and beta tests with the target population in a general classroom, ultimately converging on the SAR system used in the data collection described here.

We performed a data collection with 31 child participants aged three (2 participants) to four (29 participants) years old at the Child Development Center of East Los Angeles College (ELAC). ELAC is part of the Head Start program, so participants came from underserved and underrepresented communities. The participants were 73% Hispanic or Latino, 13% Chinese or Chinese-American, 10% Black or African-American, and 3% mixed ethnicities. 26% of the children's parents were under 25 years of age, 61% were under 31 years of age, and 55% of parents had completed some college or vocational schooling. We did not control for skill level or special needs as we intend to use the data collected to train a personalized model of number concepts learning that is viable in a general education classroom. The participants come from both Spanish-speaking and English-speaking families.

Each child participant was invited to play games with the robot one-on-one. All child participants completed an interactive tutorial on dragging and dropping virtual objects on the touch screen. In this way, we aimed to ensure that the child participants were only challenged by the learning content, not the platform, interface, or system.

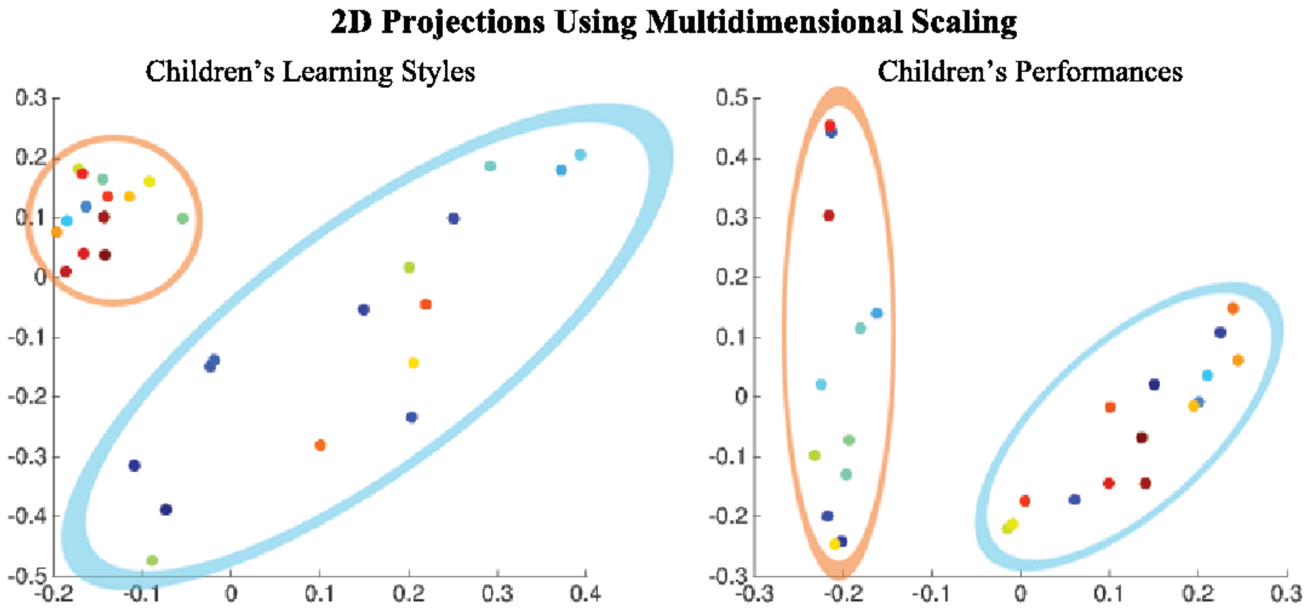


Fig. 5: Two dimensional projections of children's performance wrt learning styles (left) and wrt exercises' types (right) using kernel multidimensional scaling. Child participants cluster similarly between learning styles and exercise types.

Each exercise was randomly sampled from the same curated set of exercises for each child participant. The set of exercises spanned all seven types at various levels of difficulty. Each exercise took approximately 20-60 seconds for the child participant to complete, depending on the type of exercise and its difficulty to the child. For each child, we collected approximately 20 minutes total of child-robot interaction.

We designed and conducted this data collection as authentic to a real-world, general preschool classroom as we could achieve. We collected a substantial amount of interaction for each child participant with minimal instructor interference and a random set of exercises uniformly distributed over type and number/difficulty. Looking forward, we intend to use the collected data to generate a personalized model of number concept learning, observing both multimodal aspects of an individual child's interaction and learning style.

IV. Preliminary Analysis

After implementing the SAR system and conducting an initial data collection, we performed a preliminary data analysis to validate the data's usefulness to future personalized modeling. While not the primary contributions of this work, we found some promising patterns connecting children's interaction behaviors and learning styles to exercise performance, warranting further exploration.

On average, the child participants completed 24 exercises per session, uniformly distributed over exercise types and difficulty. The child participants had three chances to answer an exercise correctly (denoted as 1-3 in Figure 4); after three failed attempts the robot moved on to the next exercise. 67% of all the exercises performed by all participants were answered correctly on the first attempt, 11% on second, 8% on third, and 13% moved onto next exercise. While the average performance (defined by the number of attempts at an exercise) is fairly consistent across exercise types, there is a bimodal distribution over number concept (see Figure 4). This performance distribution indicates that an adaptive or personalized model of number concept learning would be exercised and useful to this target population of children. It may also be hypothesized that the number concept in an exercise, rather than the presentation of type of exercise, is the more prominent factor in determining performance.

We conducted some preliminary exploration of nonverbal behavior as an indicator of performance. As the focus of this paper is on the design process, we limit our exploration to a single modality: the child's touch interaction with the tablet. We analyzed both the child's response time, defined by the time elapsed between end of robot instruction, and the child's first move and completion time, defined by the time elapsed between the start and completion of an exercise. We found a statistically significant correlation between performance and response time ($p < 0.029$) or completion time ($p < 0.0001$). While we conclude that completion time is trivially significant as children were given multiple chances to answer correctly, the significance of response time suggests that we may be able to predict the outcome of the exercise before its completion, giving the robot the opportunity to intelligently intervene.

We conducted a further exploration into the connection between a child's learning style (defined by the parent and teacher learning style inventory questionnaire) and exercise performance. While we found that there were no statistically significant correlations between learning style and exercise performance, we used kernel multidimensional scaling to visually observe possible latent structures in a two-dimensional projection of the multidimensional learning styles and exercises data in Figure 5. Children cluster similarly and form two global structures in relation to learning style (Figure 5 left) and performance on exercises by type (Figure 5 right). This results suggests that there may be significant correlations given a larger data collection and that learning styles may be used as a prior to predict how an individual child might perform on certain exercise types.

Through our preliminary analyses of performance, learning styles, and nonverbal behavior, we have found that the number concept of an exercise is the primary indicator of the exercise's difficulty to an individual child, followed by the child's response time and learning style. These findings are supported by the literature that firstly asserts the importance of the complexity of the skill in an exercise, as defined by the educational practices and learning theory in the domain, followed by learners' affect and interaction throughout the learning process, followed controversially by learning styles. Most importantly, our preliminary analysis shows that the data collected will be useful in future training of personalized models of individual children's number concepts learning.

V. Discussion & Future Work

We have presented an interdisciplinary design process, a multimodal data collection, and initial results for a SAR approach to personalized education, specific to preschool number concept learning. We have situated our design process and approach in an interdisciplinary literature review on SAR and human-robot interaction for education, number concept learning, ML for user adaptation, ITS, multimodal computation, and learning styles theory. Through this process, we have developed a SAR approach that balances the demands of the domain–portability to and usability in a general preschool classroom, developmentally appropriate, engaging via learning-through-play, and accessible to children with limited exposure to technology–with the needs and limitations of the technology–computational characterization of number concepts and learning styles, multimodal data collection, and perceptual noise in general education classroom. Our preliminary analysis shows that the exercises were feasible for our target population, but also that a personalized model of learning would be exercised and useful and could benefit from observing the individual child's performance, nonverbal behavior, and learning styles.

In future work, we intend to explore the multimodal aspects of the children's interactions in greater detail. Using the features most significant to an individual child's performance, we plan to train a personalized number concept learning model constructed using literature-based and expert knowledge on number concepts learning. This model will then be deployed in a large user study to evaluate its ability to personalize to child participants' learning differences and the benefits of that personalization. Such multidimensionally personalized SAR systems could greatly support the efforts of educators to increase the accessibility and prevalence of interactive, socially-mediated, and differentiated education.

Acknowledgment

This work is supported by the NSF Expeditions in Computing for "Socially Assistive Robotics" (IIS-1139148).

References

[1] M. J. Mataric, "Socially assistive robotics," *Science*, vol. 3, no. 6, pp. 233–242, 1999.

[2] D. Feil-Seifer and M. J. Mataric, "Defining socially assistive robotics," in *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*. IEEE, 2005, pp. 465–468.

[3] J. J. Kuenzi, "Science, technology, engineering, and mathematics (stem) education: Background, federal policy, and legislative action," 2008.

[4] F. A. Campbell, C. T. Ramey, E. Pungello, J. Sparling, and S. Miller-Johnson, "Early childhood education: Young adult outcomes from the abecedarian project," *Applied Developmental Science*, vol. 6, no. 1, pp. 42–57, 2002.

[5] C. T. Cross, T. A. Woods, and H. Schweingruber, "Mathematics learning in early childhood: Paths toward excellence and equity." *National Academies Press*, 2009.

[6] B. Schaeffer, V. H. Eggleston, and J. L. Scott, "Number development in young children," *Cognitive Psychology*, vol. 6, no. 3, pp. 357–379, 1974.

[7] K. Schulze, R. Shelby, D. Treacy, and M. Wintersgill, "Andes: A coached learning environment for classical newtonian physics1."

[8] F. Shah, "Recognizing and responding to student plans in an intelligent tutoring system: Circsim-tutor," Ph.D. dissertation, Citeseer, 1997.

[9] P. Suraweera and A. Mitrovic, "Kermit: a constraint-based tutor for database modeling," in *Intelligent Tutoring Systems*. Springer, 2002, pp. 377–387.

[10] S. Schiaffino, P. Garcia, and A. Amandi, "eteacher: Providing personalized assistance to e-learning students," *Computers & Education*, vol. 51, no. 4, pp. 1744–1754, 2008.

[11] D. Feil-Seifer and M. J. Mataric, "Socially assistive robotics," *Robotics & Automation Magazine, IEEE*, vol. 18, no. 1, pp. 24–31, 2011.

[12] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, "The physical presence of a robot tutor increases cognitive learning gains," in *34th Annual Conference of the Cognitive Science Society (Cog. Sci.12)*, 2010.

[13] W. A. Bainbridge, J. Hart, E. S. Kim, and B. Scassellati, "The effect of presence on human-robot interaction," in *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on*. IEEE, 2008, pp. 701–706.

[14] E. Short, K. Swift-Spong, J. Greczek, A. Ramachandran, A. Litoiu, E. C. Grigore, D. Feil-Seifer, S. Shuster, J. J. Lee, S. Huang *et al.*, "How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play."

[15] N. González, L. C. Moll, and C. Amanti, *Funds of knowledge: Theorizing practices in households, communities, and classrooms*. Routledge, 2013.

[16] D. G. Singer, R. M. Golinkoff, and K. Hirsh-Pasek, *Play= Learning: How play motivates and enhances children's cognitive and social-emotional growth*. Oxford University Press, 2006.

[17] T. Kanda, M. Shimada, and S. Koizumi, "Children learning with a social robot," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 2012, pp. 351–358.

[18] D. Leyzberg, S. Spaulding, and B. Scassellati, "Personalizing robot tutors to individuals' learning differences," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 2014, pp. 423–430.

[19] J. Greczek, E. Kaszubski, A. Atrash, and M. Matarić, "Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders," 2010.

[20] S. Cassidy*, "Learning styles: An overview of theories, models, and measures," *Educational Psychology*, vol. 24, no. 4, pp. 419–444, 2004.

[21] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Engineering education*, vol. 78, no. 7, pp. 674–681, 1988.

[22] *Transitional Kindergarten Math Pacing Guide*, Paramount Unified School District, 2013.