

# Social Emotional Learning with Conversational Agents: Reviewing Current Designs and Probing Parents' Ideas for Future Ones

YUE FU, Information School, University of Washington, US

REBECCA MICHELSON, Human-Centered Design and Engineering, University of Washington, US

YIFAN LIN, Human-Centered Design and Engineering, University of Washington, US

LYNN K NGUYEN, School of Computer Science and Engineering, University of Washington, US

TALA JUNE TAYEBI, Information School, University of Washington, US

ALEXIS HINIKER, Information School, University of Washington, US

Social emotional skills are foundational competencies upon which children draw throughout their lives. This work investigates current, commercially available experiences for social emotional learning (SEL) through conversational agents (CAs). Specifically, we reviewed 3,767 Skills available in the “Kids” category of the Alexa Skills Marketplace and found 42 working Skills with connections to SEL. We found that the most common scenarios these Skills sought to support were: active listening, emotional wellbeing, conversation with other people, and politeness. The interaction patterns used by these Skills distilled into a taxonomy of styles we labeled: *The Delegator*, *The Lecturer*, *The Bulldozer*, and *The One-Track Mind*. We found that, collectively, these Skills provide shallow experiences and lack contingent feedback. To examine the gap between current offerings and families’ needs, we also conducted 26 interviews with parents to probe parents’ ideas about CAs supporting children’s SEL. Parents see potential for CAs to support children in four concrete ways, including attuning to others, cultivating curiosity, reinforcing politeness, and developing emotional awareness. Despite their optimism about these opportunities, parents expressed skepticism about CAs’ impoverished conversational abilities and worry about CAs advancing values and behavioral norms that are at odds with their own.

CCS Concepts: • **Human-Centered Computing**; • **Human-Computer Interaction (HCI)**;

Additional Key Words and Phrases: Conversational agents, social emotional learning

## ACM Reference Format:

Yue Fu, Rebecca Michelson, Yifan Lin, Lynn K Nguyen, Tala June Tayebi, and Alexis Hiniker. 2022. Social Emotional Learning with Conversational Agents: Reviewing Current Designs and Probing Parents’ Ideas for Future Ones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 52 (June 2022), 23 pages. <https://doi.org/10.1145/3534622>

## 1 INTRODUCTION

Media and interactive technology can be effective tools for social emotional learning (SEL). Children can acquire prosocial habits from television programming [37], learn to regulate their emotions through interactions with tangibles [48], and cultivate empathy through games [3]. Social emotional skills are essential to children’s long term wellbeing and interpersonal relationships [12], making interventions that support their development very

---

Authors’ addresses: [Yue Fu](#), [chrisfu@uw.edu](mailto:chrisfu@uw.edu), Information School, University of Washington, Seattle, Washington, US, ; [Rebecca Michelson](#), Human-Centered Design and Engineering, University of Washington, Seattle, Washington, US, ; [Yifan Lin](#), Human-Centered Design and Engineering, University of Washington, Seattle, Washington, US, ; [Lynn K Nguyen](#), School of Computer Science and Engineering, University of Washington, Seattle, Washington, US, ; [Tala June Tayebi](#), Information School, University of Washington, Seattle, Washington, US, ; [Alexis Hiniker](#), Information School, University of Washington, Seattle, Washington, US, .

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2022 Copyright held by the owner/author(s).

2474-9567/2022/6-ART52

<https://doi.org/10.1145/3534622>

valuable. Prior work in child-computer interaction has called for greater research into systems and experiences to support this dimension of children's growth [45].

Conversational agents (CAs) and other personified interfaces can be effective tools for SEL [?]. For example, robots and other digital agents can help children learn to resolve conflicts constructively [43] and manage their moods [44]. People respond emotionally to personified interfaces [7], and adding human-like characteristics to a system increases social and emotional reactions from users [40], suggesting CAs and other personified interfaces may be particularly well-suited to provide practice opportunities for social emotional skills. CAs are increasingly accessible to children, as smartspeakers like the Amazon Echo and mobile digital assistants like Siri have become mainstream, creating a potential pathway for delivering novel SEL supports to millions of children.

To what extent do current commercial offerings support this scenario? Our work examines, first, the prevalence of existing CA experiences that aim to support children's SEL. Further, how aligned are existing experiences with families' needs? Specifically we ask:

- **RQ1:** Do existing Alexa Skills<sup>1</sup> aim to support children's SEL?
- **RQ2:** If so, what design affordances are common across these Skills?
- **RQ3:** How do parents feel about these designs?

By examining both: 1) the current design space of CAs for SEL, and 2) parents' preferences regarding designs for this space, we sought to understand the extent to which current systems meet families' needs and to identify ways in which future work might close the gap between the two.

To investigate these questions, we first scraped metadata for the 3,767 Skills for children on Amazon's Alexa Skills marketplace, and we identified the subset that explicitly sought to support competencies related to SEL. We systematically tested these Skills to characterize their properties. We then conducted 26 interviews with parents of children between the ages of four and eight to examine parents' perspectives on using CAs for SEL generally and their perspectives on current designs specifically. We used the Skills from our marketplace review to create four exemplar storyboards that capture common user scenarios in existing Skills, and we used these storyboards as prompts during parent interviews to spark conversation.

We found only 42 working Skills that seek to support competencies related to children's SEL. We found that the most common scenarios these Skills seek to support are: active listening, emotional wellbeing, conversation with other people, and politeness. We also found that these Skills employed four main styles of interaction which we labeled, *The Delegator*, *The Lecturer*, *The Bulldozer*, and *The One-Track Mind*. Notably, we found that many Skills provided shallow experiences with minimal interactivity, and for example, 40% of Skills provided either non-contingent responses agnostic to the user's input or no response to the user's statements at all. This lack of responsiveness may limit the potential for children to have meaningful learning experiences with these Skills, as meaningful feedback is a critical dimension of SEL [13].

We further found that parents see potential for CAs to support children in four concrete ways that connect to SEL, specifically: attuning to others, cultivating curiosity, reinforcing politeness, and developing emotional awareness. Despite their optimism about these opportunities, parents also explained that they foresee design barriers to creating effective experiences in this context, including CAs' impoverished conversational abilities and the potential for CAs to advance values and behavioral norms that are at odds with those of parents.

## 2 RELATED WORK

### 2.1 Social Emotional Learning (SEL)

Social emotional skills are a foundational set of competencies for understanding the self and others that people draw on to engage in other daily activities. Children's social emotional learning (SEL) occurs through a specific

<sup>1</sup>Alexa Skills [1] are applications users can invoke and interact with Amazon Echo & Alexa Devices. Third-party developers can use Amazon's developer environment [2] to create experiences for the Echo, much like app developers create apps for mobile devices.

set of developmental tasks that are grouped under three umbrellas: 1) maintaining positive engagement in the physical and social environment, 2) managing emotional arousal and cognitive demands, and 3) maintaining positive social interaction with peers and adults [9]. One widely used framework operationalizes the competencies that children develop through SEL, labeling these: 1) self-awareness 2) self-management 3) social-awareness 4) relationship skills, and 5) responsible decision making [5].

Social emotional competencies are among the most important predictors of early school success and growth [60]. One meta-analysis shows that, for students from kindergarten through high school, acquisition of social emotional skills can contribute to better academic performance, increase positive attitudes and behaviors, enhance adjustment, and reduce emotional distress [11]. In the long term, greater social emotional competence positively affects people's educational and economic attainment, mental health, and sexual health, and interventions to support social emotional development can have effects that persist 15 years after the intervention is removed [20]. Thus, supporting children's social emotional skills is critical to children's overall development, and effective SEL interventions promise to have a meaningful impact on children's long-term wellbeing.

SEL in education is a mature field with numerous well-researched and evidence-based approaches [11]. SEL curricula emphasize experiential, active approaches, using, for example, role-play, modeling, rehearsals and practice, and simulation to support learning [47, 50]. Collectively, this work demonstrates the importance of SEL and the potential for designed supports to encourage it. Here, we build on this foundation by examining CAs as a possible site of intervention.

## 2.2 Technology that Supporting SEL

Prior work shows that media and other interactive technologies can support children's SEL [50]. Some applications designed to support the social emotional skills of older children or adults have been appropriated by young children [18], but in general, very few technologies for SEL have been developed with young children in mind [18, 38]. Of the few available technologies aiming at supporting children's SEL, nearly all specifically target children with autism [15, 16, 27, 33, 39, 54, 55]. Further, Stangl et al. note there are even fewer technologies within the Interaction Design for Children (IDC) community aiming to supporting SEL for preschoolers [50], despite the fact that this is a critical development stage for building these competencies. Collectively, this prior work suggests there is an opportunity to create SEL systems for technology's youngest users.

In analog contexts, existing evidence-based SEL interventions predominantly rely on experiential, non-cognitive learning techniques, such as role-play and modeling, which poses potential barriers to translating effective interventions to a digital context [47]. Key challenges to creating effective SEL learning experiences include, *"the lack of support for transfer and 'embedding' of skills from the SEL lessons into students' interaction, encouraging parental involvement, as well as enhancing support for development of reflective abilities and novel environments for practice"* [46]. However, other prior work explains that these barriers are not insurmountable and demonstrates that technology has the potential to contribute to children's SEL [18, 46, 47]. Specifically, researchers demonstrate that technology can enhance out-of-session learning, scaffold parental engagement, and provide feedback to curriculum developers [47].

## 2.3 Learning from CAs

Prior work consistently shows that both adults and children are able to learn from CAs [4, 26, 26, 34, 51]. For example, CAs have been shown successfully to engage children in academic activities like book-reading [59], support them in acquiring new conversational habits [22], and help them understand machine learning concepts [28].

Researchers have documented the mechanisms through which children interact with CAs, including exploration and information seeking, which contribute to their learning [31]. Researchers have found children learn more and

report higher satisfaction when learning from programs with virtual tutors due to CAs' conversational ability [57]. And other research has shown that children are responsive to CAs from an early age, and these spoken interactions can influence their early social behavior [42]. Together, this prior work demonstrates the many ways in which CAs can create effective learning experiences, even for very young children.

However, it is less clear whether CAs can support children's SEL, what commercial experiences have been created for this scenario, or how such interventions should be designed. The limited prior work on this topic has examined whether social robots and CAs can support neurodiverse individuals [25, 36, 58], and for example, Tanaka et al. found multimodal CA-based social skills training can be useful to people aging from 22 to 26 who experience social difficulties [52]. Other research has found CAs can support adolescents with Autism Spectrum Disorders in regulating negative emotions and practicing conversational skills [6].

We build on this discussion by reviewing the existing CA interactions available on the Amazon Alexa Skills marketplace that support SEL-relevant interactions, and we examine parents' perspectives on these designs as formative work for understanding the design opportunities in this space.

## 2.4 Families' Perspectives on Children's Technology

Researchers have found that parents influence children's use of technology, and parents are concerned about activities that children are engaged in with technology [8, 21, 30, 35]. Parents mediate children's interactions with technology by actively discussing it with them (active mediation), setting limits (restrictive mediation), and using it together with them (co-engagement) [41].

Other work has explored how parents and children work to co-mediate family technology use [49]. Thus, children's use of technology is widely influenced by their parents. Parents' perspective of a certain technology can determine whether their child has access to it, when and where they have access to it, and how long they access to it. Our research aims to understand parents' perspective on using CAs to support children's SEL.

Regarding CAs, Garg et al. have shown that parents want to participate in children's learning. Also, they note parents consider improving children's social behavior to be a very important learning activity [17]. However, parents are also concerned about their children using CAs. Horned notes that parents report concerns regarding their lack of control over their children's interactions with CAs, because CAs usually do not require a password to activate [24]. Further, parents are concerned about children's data being collected by CAs. Half of the parents in Garg's study mentioned that due to their children being young, they might not understand the potential harm such technology could cause [17]. Another study shows parents are concerned that interaction with CAs would influence their children's social awareness [42]. Our research builds upon these prior findings by trying to understand whether parents' perspectives are a barrier to using CAs to support children's SEL.

Other work has examined children's perspectives on CAs directly. These prior studies have found, for example, that most children enjoy interacting with CAs and typically find them to be friendly and trustworthy [10, 31, 32]. Hoffman et al. showed that children develop emotional ties with CAs, and parents reported younger children are more likely to personify CAs and believe CAs were real than older children [23]. Children enjoy interacting with CAs, especially when CAs use a friendly voice and tone, and are responsive to children [10].

Collectively, this work shows that with thoughtful design, children are able to have rich, enjoyable interactions with CA systems and learn from these experiences. Our work contributes to this discussion by analyzing what today's marketplace has made available children, whether these experience resonate with parents, and what designers might consider when entering this space.

## 3 METHOD

We conducted a two-part study composed, first, of interviews with parents of young children to explore their reaction to the idea of using CAs to support SEL in general and to specific design concepts that are prevalent

Table 1. Participant Demographics. Age, gender, and race are reported for the child; the highest level of educational attainment is reported for the parent who was interviewed. The languages listed are those spoken in the home.

PID	Age	Gender	Race	Family CAs	Languages	Highest Level of Education
P1-1	8	Female	Black	Siri and Alexa	English	Some college
P1-2	8	Female	Black	Cortana, Alexa	English, Spanish	Some college
P1-3	5	Male	Black	Alexa, Siri, Google Assistant, Cortana	English	Bachelor's degree
P1-4	7	Male	Black	Alexa	English	Bachelor's degree
P1-5	4	Male	Black	Alexa, Google Assistant	English, Swahili, Luo	Some college
P1-6	4	Male	Native Hawaiian or Other Pacific Islander	None	English, Samoan	Some college
P1-7	5	Male	Black	Alexa, Siri, Cortana	English	Associate degree
P1-8	8	Female	Black	None	English, French	Master's degree
P1-9	5	Female	Black	Siri, Google Assistant	English, Kinyarwanda	Master's degree
P2-1	6	Male	White	Siri, Google Assistant	English	Master's degree
P2-2	8	Male	White	Alexa	English, Dutch	Bachelor's degree
P2-3	8	Female	Asian, White	Google Assistant	English	Bachelor's degree
P2-4	4.5	Male	White	None	English	Associate degree
P2-5	4	Male	White	Siri	English, Russian	Master's degree
P2-6	6	Male	White	Siri, Alexa	English	Bachelor's degree
P2-8	6	Male	White	None	English	Master's degree
P2-9	8	Male	White	Siri	English	No Data
P2-10	7	Female	White	Alexa, Siri	English	Bachelor's degree
P2-11	5	Male	White	Google Assistant	English	Associate degree
P2-12	7	Male	Asian	Siri	English	Master's degree
P2-13	8	Male	White	Alexa, Siri	English	Master's degree
P2-14	6	Female	White	Siri, Cortana	English	Bachelor's degree
P2-15	5	Female	White	Alexa, Siri	English	Bachelor's degree
P2-16	6	Male	White	Alexa, Siri	English	Some college
P2-17	7	Female	White	Alexa, Siri	English	Bachelor's degree
P2-18	7	Male	White	None	English	High school degree

today. Second, we conducted a structured content analysis of all 3,767 Skills in the “Kids” section of the Amazon Alexa Skills Marketplace to examine the prevalence and design of current Skills to support SEL.

### 3.1 Interview Participants

We conducted two rounds of interviews between March and May of 2021 with 26 participants total (nine in the first round and 17 in the second, see Table 1). Participants were recruited from a database of families from the larger metropolitan area surrounding our institution who are interested in research. All participants were parents of at least one child between the ages of four and eight years old (inclusive), and in our recruiting materials, we explained that we were interested in recruiting families that use CAs in their home (though this was not an explicit inclusion criterion). In the screener that was sent to potential participants, we asked if they own any digital assistants, however, we did not exclude participants that answered no.



During the first round of interviews, we exclusively recruited families (P1-1 to P1-9) that identify as BIPOC (Black, Indigenous, People of Color), as research in HCI in the United States has historically underrepresented the views of racial minorities [29]. In the second round of interviews, we removed this constraint and expanded our recruitment to include families of any race or ethnicity. We sought to recruit people with a diverse mix of perspectives, but this sample was not intended to be representative of any particular population. Thus, we sought to elicit insights with transferability rather than generalizability [56] from this small, qualitative study.

Of the 26 participants interviewed, 19 said they currently use CAs in their homes. We asked parents to identify a target child from their household about whom they answered interview questions. On average, target children were 6.25 years ( $sd = 1.39$ ), and they included nine girls and 17 boys. After each interview, participants each received a \$30 gift card. The project was approved by our Institutional Review Board.

### 3.2 Interview Apparatus

To create the protocol, each member of the research team first individually brainstormed 8–12 potential interview questions that might help elicit responses relevant to our research questions. Then, as a group, the team discussed and consolidated these potential questions and created a first draft of the protocol.

The team then refined this draft protocol through several rounds of weekly iteration. As part of this process, the team decided to show a set of storyboards depicting existing interventions (Alexa Skills) to participants to anchor the conversation. Storyboards were intended to give parents concrete examples that might help draw out reactions, both positive and negative, and prompt new design ideas for Alexa Skills. To create these storyboards, the research team conducted a preliminary review of the Kids' section of the Alexa Skills Marketplace (this was separate from and came before the structured and complete review described in Section 3.5). Team members divided these Skills and individually reviewed their titles and descriptions to find examples of Skills that might provide an opportunity for SEL or to practice socioemotional skills. We included Skills that related to emotions, introspection, self-awareness, social communication with others, or other aspects of interpersonal relationships.

We identified 27 Skills that had a description suggesting they might connect to SEL. We then divided these Skills and research team members tried them using an Echo device. We took notes on these interactions and discussed each as a group. Through an iterative coding process using the Miro board tool<sup>2</sup>, we identified common themes from the goals of these Skills and clustered them by goal. For each cluster, we chose one representative Skill and used it as a starting point for designing a category exemplar. We selected Skills that illustrated the essence of the category and that were of comparably high quality.

We then created a storyboard to illustrate each exemplar using the Miro tool (see Figure 1). Storyboards were modeled on four existing Skills: 1) What's it Like<sup>3</sup> 2) Good Table Manners Made Easy<sup>4</sup> 3) Listening Quiz<sup>5</sup> 4) Chit Chat<sup>6</sup>. We intentionally showed the storyboards to parents at the end of the interview to avoid priming participants with examples of existing Skills before they answered earlier questions probing their general perspectives.

The final interview protocol contained five sections. We first asked participants about their child's experience with CAs, including questions like, "*How often does your child use [name of digital assistant]?*" and, "*How does your child feel about using the digital assistant?*" We then asked about the child's communication habits, social behaviors, and conversational skills, including probing the skills and habits parents value and hope their child will cultivate. Third, we asked parents about the idea of using CAs for SEL, including questions like, "*Imagine using a digital assistant to teach speech patterns to your child. What concerns do you have about this? What do you find exciting about this?*" Finally, we showed participants each of the four storyboards and asked follow-up

<sup>2</sup><https://miro.com/>

<sup>3</sup>What's it Like Alexa Skill

<sup>4</sup>Good Table Manners Made Easy Alexa Skill

<sup>5</sup>Listening Quiz

<sup>6</sup>Chit Chat

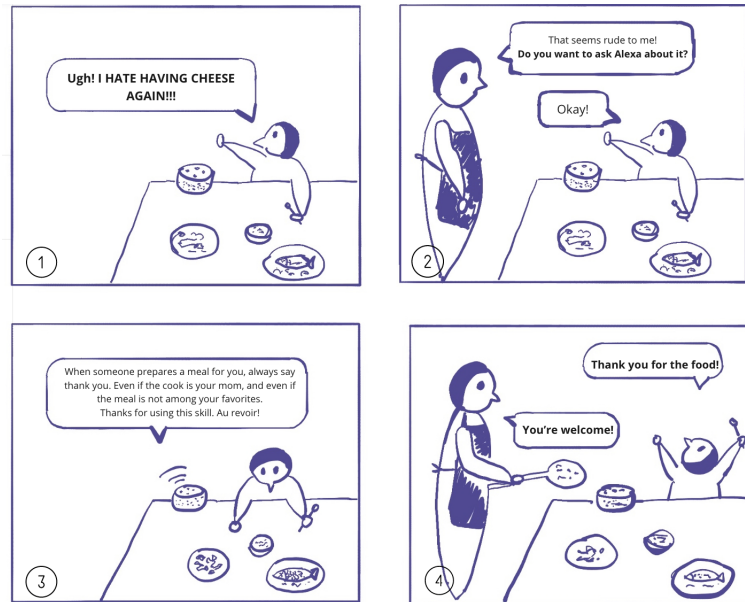


Fig. 1. Exemplar storyboard illustrating Alexa interaction

questions about each, such as, “What, if anything, do you like about this storyboard?” and “What other ideas do you have about how an Echo could support your child in learning [goal], and what do you think that would be like?”

### 3.3 Procedure

First, we emailed all of the families included in the database with background information about our research. We then sent a screener survey to those who responded with interest. After receiving consent over email, members of the research team independently led one-on-one interviews over Zoom. Interviews were semi-structured and followed the participant’s lead. The interviews were designed to take 30 to 45 minutes, and the average interview length was 39 minutes ( $sd = 11.2$ ). All interviews were audio-recorded and transcribed.

### 3.4 Interview Data Analysis

To analyze interview data, transcripts were divided across six team members. Researchers individually reviewed transcripts to identify noteworthy quotes and developed open codes. We then discussed and clarified codes, reread transcripts with the existing codes in mind, and extracted supporting and contradictory quotes using the qualitative software tool, Delve<sup>7</sup>. After multiple rounds of collaborative discussion and further refining codes, the team revisited transcripts and gathered a set of examples for each major theme. One researcher then condensed and collapsed themes to create a final codebook of interview themes and apply it across all interviews.

### 3.5 Alexa Skills Data Analysis

After completing all interviews, we revisited the Alexa Skills Marketplace to conduct a more comprehensive review of existing Skills. To determine what Skills to evaluate in depth, five members of the research team first independently coded 100 Alexa Skills, categorizing them as either relevant or not relevant to SEL. The team met

<sup>7</sup>Delve

weekly to refine the definition of what to include and discuss discrepancies and edge cases. To select these 100 skills, we chose examples to capture the diversity of existing content, sampling across all subcategories of 1) Education and Reference 2) Games 3) Music and Audio 4) Novelty and Humor 5) Health and Wellness 6) Holidays 7) Stories. The 100 sample Skills were proportionally representative of the existing distribution across these subcategories. We picked the first three skills from the first and middle page, and the last three skills from the last page for each category, skipping and replacing any Skill that duplicated a Skill that was already represented in the sample.

The research team met to collaboratively discuss their independent coding and refine the definition of what to include. Using this definition, every researcher coded a second set of 100 Skills to practice applying this definition and assess interrater reliability. We discussed disagreements until we achieved consensus and made final revisions to our definition of what to include.

We wrote a web scraper in Python using the Scrapy Library<sup>8</sup> that extracted metadata about all Skills in the “Kids” category of the Amazon Alexa Skills Marketplace. For each Skill, the scraper saved the Skill name, URL, description, number of reviews, and overall rating. This resulted in 3,767 Skills which we divided across the research team. Reviewing this scraped data, we determined that 3,279 Skills did not fit our definition of being relevant and 488 did. Of these 488, the overwhelming majority were duplicate Skills copied from blueprints. Blueprints are templates for quickly creating Alexa Skills<sup>9</sup>. We included one copy for each collection of duplicate Skills, yielding 71 total Skills. Of these, 23 could not be opened, 6 could be opened but have no following interaction, leaving the remaining data set of 42 Skills.

We then divided the Skills among members of the research team and exhaustively used each one. We took notes about the usage experience and discussed in weekly meetings. We used these session to define a coding scheme to capture interaction style, recovery style, goal, and level of contingency which we applied to all Skills.

## 4 RESULTS PART 1: EXISTING SKILLS FOR SEL

We found, first, that very few Skills explicitly sought to support children’s social emotional development. Of the 3,767 Skills we reviewed, we first identified 488 that matched our inclusion criteria based on their description and metadata alone. Of these, 379 were created with an Alexa Skills Blueprint and behaved in a manner that was nearly identical to the Blueprint (for example, offering to help people negotiate turn-taking by selecting a name at random). We tested the remaining 71 skills; 23 could not be opened, and another six could be opened but were broken such that they provided no interaction at all. Thus, we found 42 usable Skills for children that attempted to support interpersonal abilities or self-awareness in some way and might therefore foster SEL.

### 4.1 Themes in the Goals of SEL Skills

The goals of these 42 Skills clustered into four broad categories, namely: emotional wellbeing, politeness, active listening, and sparking conversation. Skills with a goal of fostering *emotional wellbeing* sought to help children label emotions, calm down when their emotions run high, build self-confidence, or otherwise understand, regulate, and enhance their emotional state. For example, the “Kid Confident”<sup>10</sup> Skill prompted children to repeat six randomly selected affirmations, as demonstrated below:

<sup>8</sup>Scrapy Library

<sup>9</sup>Alexa Skills Blueprints

<sup>10</sup>Kid Confident Skill



**Alexa:** Okay, ready? Repeat after me. Here we go. I am confident.  
**User:** I am confident.  
**Alexa:** Great job. 5 more to go. I am excellent.  
**User:** I am excellent.  
**Alexa:** Great job. We are going to do 4 more. I am respectful.  
**User:** I am respectful.  
**Alexa:** Way to go. 3 more to go. I am trusted.  
**User:** I am trusted.

Skills that supported *politeness* explicitly encouraged children to adopt normative scripts and behaviors, like saying *please*, refraining from making excuses, or sharing with others. For example, the “Good Table Manners Made Easy”<sup>11</sup> Skill instructed users to employ specific patterns of etiquette during communal meals. A child could interact with the Skill by asking Alexa to “*tell me about manners*.” Alexa would then describe scripted rules of etiquette and behavioral norms, as illustrated below:

**Alexa:** When someone prepares a meal for you, always say thank you. Even if the cook is your mom, and even if the meal is not among your favorites. Thanks for using this skill. Au revoir!  
**User:** Tell me about manners.  
**Alexa:** Some place settings include a coffee cup and saucer. If you prefer to not drink coffee or hot tea, simply turn your coffee cup upside down on the saucer. This is a signal to the waitstaff you are declining coffee. Thanks for using this skill. Good luck!  
**User:** Tell me about manners.  
**Alexa:** It is impolite to make rude comments about the food being served. Your comments will hurt the chef’s feelings. Thanks for using this skill. Bon voyage!  
**User:** Tell me about manners.  
**Alexa:** Brushing or playing with your hair at the table is not appropriate. Handle your personal hygiene needs in private, away from the table. Thanks for using this skill. Cowabunga!

A third cluster of Skills sought to cultivate *active listening*. In these experiences, the Skill provided content to listen to and a prompt encouraging the child to listen attentively, an essential competency for interpersonal connection and SEL. For example, the “Listening Quiz”<sup>12</sup> Skill first asked the user to listen carefully to a descriptive statement and then posed comprehension questions:

<sup>11</sup>[Good Table Manners Made Easy Skill](#)

<sup>12</sup>[Listening Quiz Skill](#)

- Alexa:** Astronomy is the study of everything that is or was in space beyond the Earth's atmosphere. To practice reading comprehension for tests, you'll listen to a small section on a specific topic and answer questions to test your comprehension of the material. Would you like to continue?
- User:** Yes.
- Alexa:** For thousands of years, even before we had the word astronomy, humans have been looking up at the sky. As early as 1054 A.D. astronomers in China and Japan observed the explosion of the supernova which is still visible as the Crab Nebula. Then, about 100 years later, Galileo created a new powerful telescope that could see the moons of Jupiter and spots moving on the sun. Are you ready to answer some questions?
- User:** Yes.
- Alexa:** Okay. Let's test your comprehension now. What planet could Galileo see the moons of?
- User:** Jupiter.
- Alexa:** Nice job. You've answered my question correctly. Let's continue.

Finally, a fourth set of Skills sought to spark *conversation*. These Skills posed discussion prompts or encouraged back-and-forth conversation between multiple users. For example, the “Chit Chat”<sup>13</sup> Skill generated unique conversation topics and questions for the user. After being prompted, Alexa would pose a unique question to the users. Alexa then played upbeat background music while users were encouraged to discuss their answers. The following illustrates an exchange with this Skill:

- Alexa:** Welcome back to Chit Chat. Hope you enjoy today's questions. If you could create a new holiday, what would it be called and what would you celebrate?
- User:** I want to create a holiday called International Pizza Day.
- Alexa:** Do you need more time?
- User:** No.
- Alexa:** Okay, here's another question. If you designed your own playground, what kinds of things would you like to play on? Do you need more time?
- User:** No.
- Alexa:** Okay, here's another question. If animals could talk, which one would be incredibly rude? Do you need more time?
- User:** Yes.
- Alexa:** Okay.

## 4.2 Interaction Styles

Across our dataset, we encountered four different interaction styles. Although other patterns of interaction may exist, these were the only common patterns we encountered, and the Skills we examined all used at least one of these styles. Some Skills employed only one, while others drew on different styles in different parts of the experience (see Figure 3). This taxonomy of interaction styles included:

- **The Lecturer:** The device talked *at* the user, telling them things without interacting. For example, the “Daniel Learns Sharing” Skill began by telling a story. The skill continued talking without any prompt for

<sup>13</sup>Chit Chat Skill

user interaction until the end when it asked the user to choose whether to quit the Skill or to listen to the story again.

- **The Delegator:** The device tried to prompt interactions between people but did not participate itself. For example, the “Chit Chat” Skill prompted the user to converse with the people around them, asking questions like, “*What food would you like to eat in space?*” Although it posed a question, the Skill did not listen for a response. Instead, it played instrumental music for 90 seconds (during which time, the user would presumably be engaged in conversation with another person). The Skill then gave the user the option to continue talking to their conversational partner or to move on to a new discussion prompt. At no point did the Skill engage with or provide feedback in response to the user’s conversation.
- **The Bulldozer:** The device interacted with the user but continued on its conversational path no matter what the user said. For example, the Skill “Kid Confident” began by asking users to “*Repeat after me.*” Then, regardless of how the user responded, the Skill replied by saying “*Great job, we are going to do [number of remaining repeated sentences] more.*” Although the Skill prompted the user to respond (unlike The Lecturer), it did not listen for a response. The Skill “bulldozed” to the end, talking over the user and ignoring the responses it requested from the user.
- **The One-Track Mind:** The device interacted with the user but only within a very narrow script (such as forced-choice responses) on a topic of its choosing. In these instances, the Skill was able to respond, because the input space was constrained to be almost entirely unambiguous. For example, “Calm Now” asked the user to describe their feelings by saying things like, “*Thinking about how you’re feeling right now, would you say you are sad, mad, or both?*” and “*Thinking of your sadness and anger right now, are you feeling it a lot, medium, or a little?*” These interactions enabled the Echo to respond in a way that was tightly coupled with the user’s input but severely restricted the input that would be considered valid.

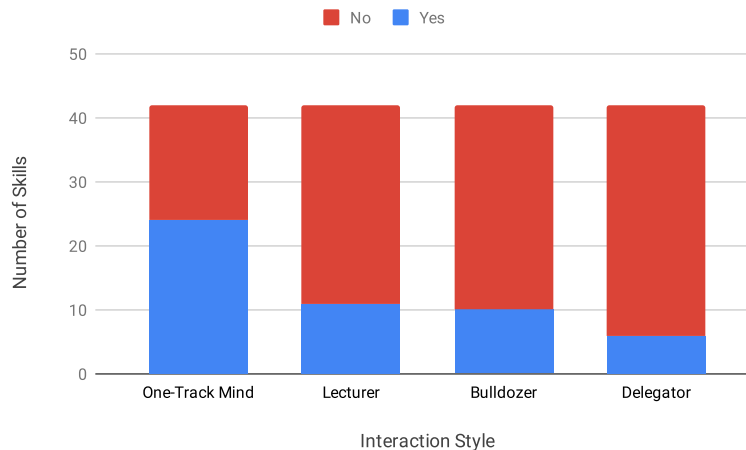


Fig. 2. Number of Skills that used each of the four interaction styles we identified.

The most common interaction style within the Skills we tested was the One-Track Mind, which appeared in 57% of Skills. The second most common was the Lecturer, which appeared in 26% of Skills. Figure 2 shows the frequencies of all four styles.

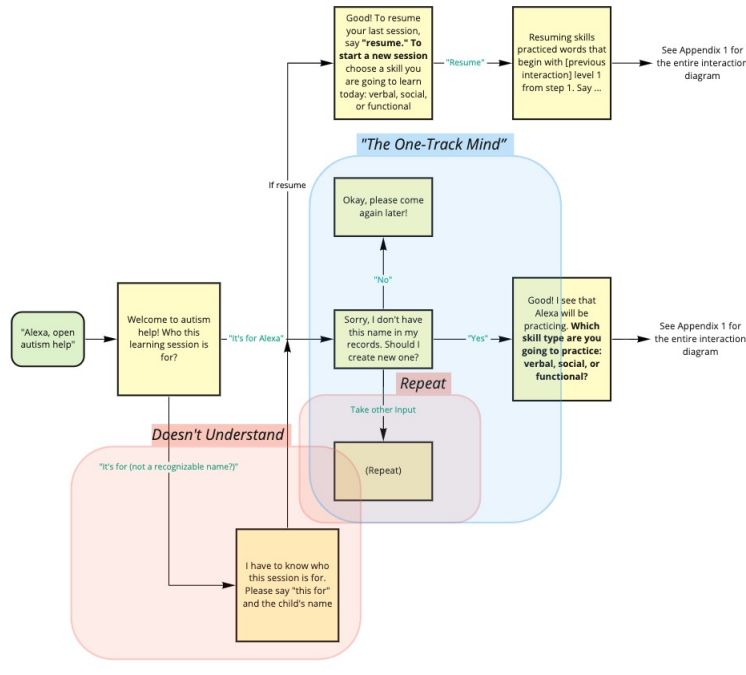


Fig. 3. A diagram of a subset of interactions from the Skill “Autism Help” (see Appendix A for the entire interaction diagram). Blue shading highlights interaction styles, and red shading highlights recovery strategies.

### 4.3 Feedback and Contingent Responses

In connection with these interaction styles, we found that many Skills provided very little contingent feedback to users. Of the 42 working Skills we tested, 40% provided feedback that was agnostic to the user’s input. For example, the “My Compliments Shine” Skill allowed for any input and always responded by giving the user a compliment:

**Alexa:** You’re perfect just the way you are. Want to know something else?  
**User:** Hamburger.  
**Alexa:** Wonderful. Want to hear another good thing?

Skills frequently provided these kinds of non-contingent responses, which we came to refer to as “failing the hamburger test.” In these instances, even if the user interjected a nonsensical term like “hamburger,” the Skill would proceed seamlessly, as if it had received valid, expected input. Only 48% of the 42 working Skills we examined passed the hamburger test during the Skill’s main user scenario.

Only four Skills (9%) attempted to respond contingently to open-ended input from users. And despite their attempts to simulate human-to-human conversational flow, these Skills displayed limited ability to offer rich, appropriate responses. For example, in the following exchange, the “Spare Parts” Skill attempted to engage the user in conversation:

**Alexa:** There is something wrong with my GPS module. I think I'm lost. Have you ever been lost?

**User:** Yes.

**Alexa:** Oh no! What did you do?

**User:** I wandered off by myself in the store and got lost from my parent.

**Alexa:** Being lost is no fun. To fix my GPS module, we need to fix my global language settings. I will change my accent and you will tell me what country I sound like I am from.

#### 4.4 Recovery Strategies

We encountered three common recovery strategies that Skills used when they could not understand the user or when an interaction failed. These strategies included:

- **Repeat:** The CA says it does not understand and repeats its prior statement. For example, when the Skill "Robot Says" did not understand what the user said, it repeated the previous statement. If it was unable to interpret the user's response after several rounds of repetition, it quit abruptly.
- **Continue:** The CA continues the interaction, ignoring the failure.
- **Break:** The CA does not respond at all and the interaction fails. For example, when the Skill "Happy Birthday in Many Languages" did not understand the user, it quit without comment or notification.

Most Skills used a combination of all three recovery strategies. For example, when the Skill "Esme & Roy" failed to understand the user, it would first say it did not understand and repeat the question. If it failed to understand a second time, it continued to the next interaction.

## 5 RESULTS PART 2: PARENTS' PERSPECTIVES ON USING SKILLS FOR SEL

### 5.1 Conversational Skills Parents Value and Opportunities to Support Them

Parents described valuing several common abilities concerning their child's interactions with other people. We categorize these competencies into five groups below, specifically: attunement, social confidence, self-regulation, curiosity, and politeness. Here, we describe how parents characterize these abilities and the opportunities they see for CAs to support them.

**5.1.1 Attunement.** Many parents described valuing children's ability to attune themselves to others in conversation. Attunement refers to a deep awareness and precise responsiveness to others [19], and Erksine describes it as "*a kinesthetic and emotional sensing of others-knowing their rhythm, affect and experience by metaphorically being in their skin, and going beyond empathy to create a two-person experience of unbroken feeling connectedness by providing a reciprocal affect and/or resonating response*" [14]. In interviews, parents described the importance of cultivating this kind of fluid, real-time awareness during conversations with others through, for example, eye contact, responsiveness, and awareness of turn-taking.

For example, some parents described their child's ability to attune to others in conversation as something they are particularly proud of, saying things like, "[he] is most observant about small details, so he would be the one who could remember what someone was wearing or that they changed their haircut" (P2-1). Many more parents described children's struggles to attune to others and proactively mentioned this as one of the primary areas in which they hoped to see their child grow. They said things like, "I think just with time he'll understand social cues a little better and he won't interrupt as much" (P1-3), "[we are] still working on him listening to people's responses and then replying to what they actually said, as opposed to just continuing his own thought process" (P2-2), and "oh my goodness yes she doesn't know when the joke's not funny anymore" (P2-17). In these instances and others,

parents described various ways in which children showed a lack of attunement, and their examples reflected challenges with turn-taking and conversational patience, difficulty picking up on subtle cues in conversation, and a general lack of dexterity and ability to nimbly shape conversations in response to the needs of others.

When asked to describe ways in which CAs might support SEL, parents frequently brought up aspirational scenarios in which a device created a context for children to practice attuning to others in conversation. For example, they said things like, *“it’d be nice if Alexa prompted communications with them and wasn’t a passive responder so that they could understand when somebody engages them in a conversation how they can develop responses”* (P2-13). They wished for an agent that could help a child practice *“how to make sure that you’re listening to what the person is saying in a conversation...just being present in a conversation”* (P1-2). Another parent described using a CA’s existing ability to tell jokes as a sandbox for practicing conversational turn-taking, saying, *“they all love it because they’re doing it while they’re packing backpacks or lunches, and they can all laugh together, and it does teach kids that turn-taking”* (P2-1). They said that they wished CAs would model attunement and respond contextually to the tone of the speaker, saying, *“if there was a certain tone or a demanding type of tone [from the user], if we could get her [the CA] to maybe not respond or respond a certain way, to get him [the child] to sort of switch it up”* (P1-4).

However, parents also frequently pointed out the inability of CAs to attune to the user as a major shortcoming of the form factor. They were skeptical that a CA could effectively support children’s SEL, because CAs themselves lack human-like sophistication to adapt fluidly in conversation and attune to their conversational partner. Parents described this inadequacy saying things like, *“she doesn’t really care in what tone of voice you’re asking the question or how you’re responding”* (P2-2) and, *“for me there’s so much more to communication than just something that could be taught by, at this time, a machine, which I don’t think it’s equipped to assess those sorts of nuances, which for me are a very human thing”* (P2-12). Thus, parents saw attunement as both an important conversational skill where children would benefit from support but also an area in which the device itself struggled. One common suggestion from parents was scaffolding for simple turn-taking interactions, which parents saw as beneficial for children’s SEL and also manageable for CAs. For example, several parents suggested things like, *“if it somehow was able to model when it’s appropriate to stop talking and allow someone else to take a turn”* (P2-10).

**5.1.2 Social Confidence.** Parents also frequently expressed the hope that their children would develop conversational confidence in social settings. Parents expressed this desire in two different ways, with some parents sharing concerns about their child’s current *lack* of social confidence and others expressing pride in their child’s already-robust social confidence. For example, parents referred to their child’s shyness in front of strangers or in public, saying things like, *“my eight-year-old, she’s very shy around people that she doesn’t know”* (P1-1), *“he’s just very shy, and so while he has high vocabulary at times, just in general that shyness kind of stifles his ability to communicate freely with people he doesn’t know or people he does know”* (P2-12), and *“My kids are very outgoing; they’re very into sports and all of that, but I’ve noticed that they’re more shy and hesitant to speak out to people...if we’re going out in public.”*(P1-4)

Parents described children demonstrating this shyness in different ways, including hiding behind parents (P1-4), declining to make eye contact (P2-6), or being unable to express their needs in front of others (P1-6). These parents also expressed their hope that their child would grow in confidence in social settings. For example, one parent described replying to her child when the child asked to use a swing at the playground; the parent reported responding with:

*“Well, just go over there and say, ‘can I please have a turn?’ And so she’s [the child] brave. She went over there and said, ‘can I please have a turn?’ And at first, they [an unknown child currently using the swing] couldn’t hear her, but she kept working on it”* (P2-15).



In these and other instances, parents explained that they hope their child will grow to self-advocate, initiate conversation with others, and have the confidence to express themselves authentically.

Parents also expressed the belief that CAs have the potential to support the development of social confidence and to help children become more comfortable communicating with peers and adults. For example, one parent explained she would appreciate CAs' support to help her child learn *"how to not be awkward when you're meeting new people or not to be shy... so just teach how to get rid of those awkward moments so that she just feels comfortable with her peers and with herself and with other people"* (P1-1). Another parent mentioned, *"I think kids can be shy sometimes so helping them come up with questions to ask other kids to become friends or to start that dialogue"* (P2-6).

**5.1.3 Curiosity.** Parents also described valuing children's curiosity and inclination to ask questions. For example, one parent said, *"I'm really proud of her curiosity. It's one of the joys of being a mom of a five-year-old, they are so curious. And they ask a bunch of questions, they want to know everything"* (P1-9). The participant continued, explaining that this curiosity can help her children, saying, *"Obviously it's good for her learning; it's good for her brain development, being curious. So I would like her to continue with that"* (P1-9). Other parents also said they appreciated their children asking questions, or wishing their child would show more curiosity, saying things like, *"some more of asking, in terms of asking [others'] opinion, or their ideas, things like that."* (P2-11), and *"he likes to ask questions, so at the end of the day, any questions he has from the day he wants them answered before he goes to bed... He's really curious and he's able to communicate that."* (P1-3).

Parents further said that they saw an opportunity for CAs to support children in cultivating and leveraging their curiosity. For the participant 1-9 mentioned above, when asked about how CAs can support children's communication, the parent mentioned, *"it's really like asking questions and how questions are appropriate at this age, just mainly curiosity question, [asking] curious questions"* (P1-9). Another parent directly explained, *"That if it [the CA] was used... in a way to satisfy their curiosity about like things you would look up back in my day at an encyclopedia or something, I think that would be exciting to them... That would be fun."* (P2-8). One parent mentioned that a CA to satisfy children's endless curiosity would reduce the burden on the parent to do so, saying, *"it's nice to have a third party to answer questions and weigh in with conversations from the endless curiosity of kids and give my wife and I a break, so it's kind of cool"* (P2-13).

**5.1.4 Politeness.** Parents also value politeness in their children's interactions. Parents frequently mentioned that they expect their children to say "please," "thank you," and other social niceties. For example, one parent explained, *"I like that he'll say 'thank you' often when I give him something. He appreciates things. He'll say sorry when he knows that I'm upset about something"* (P2-4). Other parents described their children by saying, *"I think he answers adults very politely"* (P1-3) and, *"he's usually pretty polite. I think things that he's heard he'll often—like if a grandparent says, 'Do you want a bite of my broccoli?' and it's something he doesn't want, and he'll say, 'Thank you for asking, but no, thank you'"* (P2-8).

Other parents described their children's struggles and developing ability to be polite. They explained that they wish their children were *"a little bit better about the asking nicely, not needing as many reminders to [say], 'May I please have,' or, 'Would you please do this,' as opposed to, 'I want milk'"* (P2-11). Still others described this challenge saying things like, *"[The child becomes] The center of attention. Me, Hey, me, me, me, look at me... I'm like, 'You're little. So I guess it can kind of be about you right now, but you better snap out of it because the world does not revolve around you, sunshine'"* (P1-1).

Parents' common interest in supporting developing politeness suggests an opportunity for CAs to provide a context for practicing social niceties. As one parent suggested, *"It would be great if she [Alexa] required him to say like, 'please,' and 'thank you.' Or if there was like a feature where they practice manners"* (P1-4). Other parents suggested the device provide *"pretend scenario[s] where they're practicing manners"* (2-9) and *"more instances*

during the day where, you're at home with the Echo where it could be jumping in and correcting your kids. Like making sure your kids said please and said thank you" (P1-7).

**5.1.5 Emotional Awareness.** Finally, participants told us that they value their children developing an awareness of their own emotional state. A number of parents mentioned that they value children's self-awareness and self-control, saying, *"I think patience with speaking to, because I noticed that now this generation of my kids that we're having, it has very little patience and they just don't want to wait for you to finish your sentence"* (P1-4). Other parents pointed to the importance of understanding others' emotions in conversation, for example, expressing pride that their child *"is very empathetic and can understand emotions and how people are feeling and project that their feelings onto himself to understand it"* (P2-13).

When asked to describe ways in which CAs might support children's emotional awareness, multiple parents suggested that CAs could help children to check in on their own feelings. They said things like, *"If the Echo could ask questions... and respond to the kids talking, like asking about, 'how did it feel? How did things feel?' and I think that'd be helpful"* (P1-7), and *"I think it'd be interesting if you start your day or one time a day or something like that. You kind of check in with Alexa, and Alexa's like, 'how do you feel today?'... It's almost like you check in with Alexa, and then Alexa can start to catalogue and it'll be like, 'wow, you felt good three days this week' or something like that"* (P2-15).

Parents also saw potential for CAs to scaffold the child in deepening their capacity for introspection. One parent suggested:

"I guess I hear kids a lot of times answer that question with like, 'I felt good. I felt bad. I liked it. I didn't like it,' in very simplistic terms. And I don't know if there would be a place for the assistant to kind of probe that a little further, like getting more into nuances of those feelings" (P2-18).

Similarly, other parents suggested, *"I wonder if there would be some way, if they hear something like the child saying, 'I'm angry or upset,' having something like, 'Would you like me to guide you in some way to calm down?' Or things like that. 'Should we practice breathing'"* (P2-11). Another parent expressed similar interest in scaffolding emotional self-regulation, saying, *"When she's having—we call them feelings bursts, which is when you have a big overwhelming feeling—all of a sudden... If Echo wanted to somehow walk my child through how to deescalate, how to calm down, how to breathe. Things that we do as parents in person, but perhaps we're not always there, that would be awesome"* (P2-10).

## 5.2 Parents' Concerns about Using CAs for Socioemotional Learning

However, parents' responses were far from universally positive, and many expressed ambivalence or skepticism about the idea of using CAs to support SEL. We encountered three common types of concerns, which we elaborate on below.

**5.2.1 Inadequate conversationalist.** First, parents were concerned that CAs could not possibly live up to the conversational demands of SEL scenarios. Parents worried that, due to technological and design constraints, CAs would be insufficient conversational partners for children in this context. Some parents expressed skepticism about CAs' ability to respond contingently, a concern that was validated in our analysis of marketplace Skills (see Section 4.3), where Skills lacked contingency, depth, and conversational context. One parent explained, for example, *"I believe a conversation is meant to develop and to take on a topic matter and a certain level of nuance as you navigate between the various responses and this is just eliciting a full response, but then there's no continuation or relationship between question and answer of the conversation"* (P2-13).

Other parents explained that they were skeptical of CAs' ability to be adequate conversationalists in the context of SEL because of their lack of emotional expression in their speech emotions, and lack of tone, body language,

and other implicit human-to-human communication channels. For example, one parent mentioned that *"I say 'she' because Alexa's voice is female, but she doesn't really care in what tone of voice you're asking the question or how you're responding. She only needs to hear the words clearly"* (P2-2). In addition, some parents explained that they did not believe CAs would be able to respond adequately to complexities of SEL-related contexts. As one parent said, *"I don't think this will be realistic. And I don't know if Alexa will be so complex to answer all the angles my five-year-old will have, if she was having this conversation with her"* (P1-9).

Lastly, parents expressed the concern that CAs' lack of conversational sophistication would render them unable to engage children in the context of SEL. They said, *"The problem will be having the attention of a child to watch it or listen to it"* (P1-5), and *"it doesn't seem like it's going to be capturing the engagement or imagination of the child"* (P2-2).

**5.2.2 Values Misalignment.** Many parents expressed concerns about CAs promoting behaviors and norms that might conflict with their values. They said things like, *"I guess I'm just looking at it from a parental perspective that I would prefer that Alexa agree with the parent"* (P1-6),

*"My objection to anything like that would not be the objection to the device doing the thing. It would be the objection potentially to the way the device was doing. If it was teaching her something I didn't like... I would have to evaluate what it was teaching real-time to see if I agreed or disagreed"* (P2-3), and *"that always scares me because I feel like if somebody else teaching and influencing them"* (P1-1).

The context of SEL heightened parents sensitivity to the importance of designing systems that reinforce the parents' own messaging to children. They explained, *"I just don't think that robots should be teaching kids... social interactions, because there are so many angles that it could go the wrong way"* (P1-9) and *"I think manners are also subjective. So I also think just a definitive, digital assistant answer for manners is probably not a question I would ask"* (P1-3). Parents were skeptical that CAs could be programmed to represent their values accurately and expressed concern about devices weighing in on nuanced and personal choices.

Parents suggested that the device support and supplement the parent's agenda rather than bringing its own. One parent explained:

*"I feel like being a backup for the parent rather than a step-in for the parent is a more powerful thing. It is still giving the human parent the gravity in the moment. Rather than the device. So, if the parent had a different kind of prompt they could do like, 'Hey, that wasn't an okay way to say that, maybe we should ask Alexa about it.' And Alexa, 'Like, I agree.' Backing up the parent, a united front can work rather than just no, we're going to let the assistant tell you what's going on."* (P2-17)

Other parents also suggested, for example, that a CA would be *"good as a supplement"* and useful for *"reinforcing learning."* As one parent explained, *"I like the idea that the computer or Alexa would come in as backup to mom and reinforce better behavior"* (P2-14).

**5.2.3 Replacing Attachment Relationships.** Finally, many parents raised concerns about the idea of a device taking on a task that they saw as the purview of the parent. They said things like, *"this is a job of a parent or a teacher"* (P1-9) and, *"the mom should be taking on responsibility"* (P2-13). These concerns went beyond the worry that a device would do a poor job with the task or adopt values the parent disagreed. In these instances, participants expressed concerns with the very premise of interacting with in a context that is usually reserved for close relationships. And parents saw the idea of a device taking on parenting responsibilities as problematic because of its threat to the parent-child relationship. They told us, *"I don't like the idea of an electronic device replacing the parent's relationship"* (P2-16) and explained that they want to prioritize *"having that connection with their kid versus like kind of shoveling their kid off to Alexa"* (P1-4).

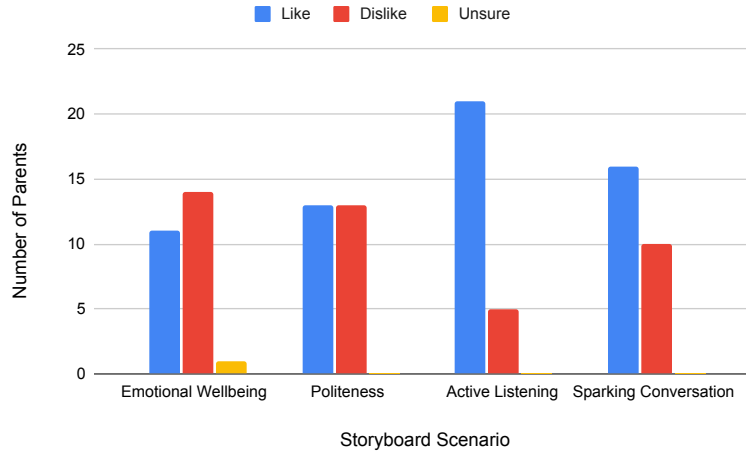


Fig. 4. Parents' reaction to storyboards

Similarly, some parents expressed concerns about the idea of CAs filling any human-to-human role and were dissatisfied with the idea of devices undermining any relationship. They said things like, “*but just the idea of getting a friendship with a device rather than a human just doesn't seem natural*” (P2-4) and, “*the children will learn by interacting with other human beings, to be honest, not with the robot*” (P1-9). To combat this challenge, several parents suggested the experience be collaborative and involve bringing people together. They offered ideas like, “*If there's good learning from it, like social emotional coaching, all three of my kids are relatively close in age, and it would be great for them all to hear that together. So maybe they're all sitting on the couch, listening to it together*” (P2-1), and “*maybe she [the parent] could have been there together with the child asking like, 'let's ask Alexa, you know about this'*” (P1-4).

### 5.3 Parents' Reactions to Storyboards

Finally, we examined parents' reactions to the four design scenarios we presented (see Figure 4). Their reactions reflected the concerns and interests described above, and for example, parents raised concerns about inadequate conversational abilities and the inappropriateness of replacing parents. The most popular storyboard portrayed an interaction in which the device encouraged active listening, which 80% of parents said they would be interested in trying at home. Parents explained that this was because they preferred CAs to give objective, deterministic answers instead of subjective opinions. They were concerned about unsupervised subjective answers CAs may give to their children. For example, one parent said, “*I like it a lot because I think it can be effective in teaching active listening or comprehension even. But the other ones, what I didn't like was because there were no concrete answers*” (P1-9).

## 6 DISCUSSION

### 6.1 CAs for SEL: The Current Designs and Future Opportunities

We found that, today, there are very few Skills in the Kids' Section of the Amazon Skills Marketplace that connect to social emotional competencies. The ones that do exist follow limited interaction patterns and employ simplistic recovery strategies that rely on repetition and are quick fail when they do not understand the user. Given

the importance of contingent feedback in SEL interventions, this minimal interactivity may be inadequate for meaningful learning gains. A “Bulldozer” may make for a poor SEL facilitator, and Skills that fail “the hamburger test” are unlikely to create the kind of authentic, meaningful exchanges that are necessary for social emotional development [19].

However, the majority of the parents in our study saw promise in the idea of designing SEL supports for CAs, even as they foresaw barriers as well. There were notable gaps between parents’ vision of what CAs might offer and what CAs offer today. For example, parents envisioned systems that would expect politeness, promote curiosity, and encourage children to practice social confidence. These goals for children’s social emotional development suggest possible design concepts for future systems. Many parents suggested CAs scaffold patient turn-taking (rather than tolerating any interruption) and expect politeness from their users.

Parents were also open to the premise of adopting such systems and suggested many design concepts of their own. They envisioned experiences that siblings would engage with together, imagined Alexa speaking up to back up the parent when they set a norm for the child, and they envisioned a check-in app that coaches a child through discussing their feelings. The vast majority of parents were excited about the idea of a design that would provide a context for practicing active listening and said that this aligned with their child’s struggles to be attentive in conversation, share respectfully, and give the speaker space to finish their thoughts without interruption. These design ideas suggest an openness to adopting interventions in this space, and they surface specific user needs and design concepts to explore.

## 6.2 CAs for SEL: Design Barriers

However, parents also expressed concerns about CAs replacing human connection, disrupting their relationship with their child, and advancing values that differ from their own. Parents hoped to see systems augment their own parenting and reflect values they set, consistent with prior recommendations in the The People + AI Research guidebook<sup>14</sup>. This prior work urges designers and technologists to gauge when intelligent systems should take unsupervised, automated action, and when augmentation would be more appropriate instead. The authors explain that automation is, “*often the best option for tasks that supplement human weaknesses with AI strengths*,” while augmentation is recommended in personal, high-stakes, or emotional circumstances and in circumstances where users have social obligations to others. Thus, CAs have promise as a technology for supporting SEL, but will be more likely to support parents’ values if they augment pre-existing processes in the home rather than attempting to replace them.

Given parents’ discomfort with CAs holding authority in social interactions, designers might, for example, enable parents to verbally cue CAs to reinforce certain norms, such as speaking politely. To bolster the focus on parent-child relationships, Skills can be designed for co-engagement (such as prompts to debrief a Skill with parent). Interactions can also incorporate built-in transparency that highlights the social limitations of a CA, with cues such as, “*ask another adult*.” Designers can also follow parents’ suggestion to design SEL interventions that have well-understood correct answers that do not require the device to wade into nuanced territory that will be better navigated by a human.

Parents’ responses also made clear that technical limitations introduce additional design barriers. Currently, state-of-the-art commercial CA devices fall short of providing desired functions and interactions parents described. They cannot offer contingent and knowledgeable answers to children’s open-ended questions, they have difficulty understanding children’s open-ended responses, they cannot remember previous interactions, and they lack sophisticated recovery strategies to draw upon when interactions fail. Parents saw these shortcomings as potentially insurmountable, and their concerns were supported by our review of current Skills, where we found that existing interventions conform to a small set of interaction styles with limited potential for fostering SEL. Thus, future

<sup>14</sup><https://pair.withgoogle.com/guidebook/>

work might consider: 1) innovating technical advances improving the attunement capabilities of the device, 2) examining interventions for SEL within a narrow range of possibilities (i.e., supports for practicing active listening and turn-taking, or story-based interventions supplemented with forced-choice questions), and 3) investigating components of SEL where CA-based interventions would be counterproductive.

### 6.3 Limitations

The most important limitation of this work is that it does not incorporate children’s perspectives directly, and we plan to build on this foundation by investigating the types of SEL experiences children themselves would (and would not) like to explore. Another limitation of this work is the small sample size and lack of male interviewees. We also examined commercial content only in a single marketplace. Expanding upon the present work to probe the perspectives of fathers and children, and to examine offerings for other systems, would likely yield additional insights. Future work also remains to explore live interventions in the home and to understand how families engage with working prototypes.

## 7 CONCLUSION

We conducted a two-part study in which we examined current Amazon Echo Skills for supporting SEL, drawing on an analysis of 3,767 existing Skills. We compared these existing designs to the perspectives and needs of parents, who reflected on how CAs might support children in this context. Parents see several ways in which CAs might support children’s SEL, such as encouraging curiosity and supporting attunement to others; these ideas only partially overlapped with the goals of existing Skills. This suggests an opportunity to create novel designs that better support families. Despite their willingness to try interventions in this space, parents also raised concerns, saying that today’s CAs are inadequate conversationalists for discussing nuanced topics and, more strikingly, that they worry about CAs intruding on territory that should be reserved for humans alone.

The current set of SEL-related Amazon Echo Skills offer shallow experiences with minimal feedback that are far from the careful designs parents envision. Skills used a small set of interaction styles with minimal interpersonal sophistication, and, for example, “lectured” and “bulldozed” their users. They consistently failed our “hamburger test,” which evaluated whether a Skill had *any* awareness of the user’s input. Taken together, these findings suggest this is a design space that is underdeveloped and has potential to provide value to families. Technical limitations may place a ceiling on CAs’ potential to support SEL, but parents still outlined many concrete ways in which today’s devices could support children. Future designs might scaffold children in taking turns without interrupting, provide support for labeling emotions, provide practice opportunities for active listening, and create a baseline expectation of politeness when making requests.

## ACKNOWLEDGMENTS

Thanks to Meena Devii Muralikumar, Shengzhi Wang, and Savanna Yee for their help with this research. This work was supported by the Jacobs Foundation and the Canadian Institute for Advanced Research.

## REFERENCES

- [1] Amazon. 2022. Alexa Skills. <https://developer.amazon.com/en-US/alexa/alexa-skills-kit>, Last accessed on 2022-04-25.
- [2] Amazon. 2022. Alexa Skills Kit SDKs. <https://developer.amazon.com/en-US/docs/alexa/sdk/alexa-skills-kit-sdks.html>, Last accessed on 2022-04-25.
- [3] Alissa N Antle, Ofir Sadka, Iulian Radu, Boxiao Gong, Victor Cheung, and Uddipana Baishya. 2019. EmotoTent: Reducing School Violence through Embodied Empathy Games. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (Boise, ID, USA) (IDC ’19). Association for Computing Machinery, New York, NY, USA, 755–760.
- [4] Cynthia Breazeal, Paul L Harris, David DeSteno, Jacqueline M Kory Westlund, Leah Dickens, and Sooyeon Jeong. 2016. Young Children Treat Robots as Informants. *Top. Cogn. Sci.* 8, 2 (April 2016), 481–491.
- [5] CASEL. 2020. CASEL’S SEL Framework. <https://casel.org/casel-sel-framework-11-2020/>, Last accessed on 2022-02-13.



- [6] Inha Cha, Sung-In Kim, Hwajung Hong, Heejeong Yoo, and Youn-Kyung Lim. 2021. Exploring the Use of a Voice-based Conversational Agent to Empower Adolescents with Autism Spectrum Disorder. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21, Article 42). Association for Computing Machinery, New York, NY, USA, 1–15.
- [7] Kate Darling, Palash Nandy, and Cynthia Breazeal. 2015. Empathic concern and the effect of stories in human-robot interaction. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 770–775.
- [8] Vauna Davis. 2012. Interconnected but underprotected? Parents' methods and motivations for information seeking on digital safety issues. *Cyberpsychol. Behav. Soc. Netw.* 15, 12 (Dec. 2012), 669–674.
- [9] Susanne A Denham, Hideko H Bassett, Melissa Mincic, Sara Kalb, Erin Way, Todd Wyatt, and Yana Segal. 2012. Social-Emotional Learning Profiles of Preschoolers' Early School Success: A Person-Centered Approach. *Learn. Individ. Differ.* 22, 2 (April 2012), 178–189.
- [10] Stefania Druga, Randi Williams, Cynthia Breazeal, and Mitchel Resnick. 2017. Hey Google is it OK if I eat you?. In *Proceedings of the 2017 Conference on Interaction Design and Children* (Stanford California USA). ACM, New York, NY, USA.
- [11] Joseph A Durlak, Roger P Weissberg, Allison B Dymnicki, Rebecca D Taylor, and Kriston B Schellinger. 2011. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Dev.* 82, 1 (Jan. 2011), 405–432.
- [12] Richard A Easterlin. 2003. Explaining happiness. *Proc. Natl. Acad. Sci. U. S. A.* 100, 19 (Sept. 2003), 11176–11183.
- [13] Maurice J Elias and Others. 2003. *Academic and social-emotional learning*. Vol. 11. International Academy of Education Brussels, Belgium.
- [14] Richard G Erskine. 1998. Attunement and involvement: Therapeutic responses to relational needs. , 235–244 pages.
- [15] Lizbeth Escobedo, David H Nguyen, Louanne Boyd, Sen Hirano, Alejandro Rangel, Daniel Garcia-Rosas, Monica Tentori, and Gillian Hayes. 2012. MOSOCO: a mobile assistive tool to support children with autism practicing social skills in real-life situations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (CHI '12). Association for Computing Machinery, New York, NY, USA, 2589–2598.
- [16] Jose Maria Garcia-Garcia, Victor M R Penichet, Maria D Lozano, and Anil Fernando. 2021. Using emotion recognition technologies to teach children with autism spectrum disorder how to identify and express emotions. *Universal Access in the Information Society* (June 2021).
- [17] Radhika Garg and Subhasree Sengupta. 2020. Conversational Technologies for In-home Learning: Using Co-Design to Understand Children's and Parents' Perspectives. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.
- [18] Naska Goagoses, Heike Winschiers-Theophilus, and Michael Bosomefi Chamunorwa. 2020. The (Potential) Role of Technology for Young Children's Social-Emotional Learning: During and Beyond the COVID-19 Pandemic. *Digit. Gov. Res. Pract.* 2, 1 (Dec. 2020), 1–6.
- [19] John M Gottman. 2011. *The Science of Trust: Emotional Attunement for Couples*. W. W. Norton & Company.
- [20] J David Hawkins, Rick Kosterman, Richard F Catalano, Karl G Hill, and Robert D Abbott. 2008. Effects of social development intervention in childhood 15 years later. *Arch. Pediatr. Adolesc. Med.* 162, 12 (Dec. 2008), 1133–1141.
- [21] Alexis Hiniker, Sarita Y Schoenebeck, and Julie A Kientz. 2016. Not at the Dinner Table: Parents' and Children's Perspectives on Family Technology Rules. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) (CSCW '16). Association for Computing Machinery, New York, NY, USA, 1376–1389.
- [22] Alexis Hiniker, Amelia Wang, Jonathan Tran, Mingrui Ray Zhang, Jenny Radesky, Kiley Sobel, and Sungsoo Ray Hong. 2021. Can conversational agents change the way children talk to people?. In *Interaction Design and Children* (Athens Greece). ACM, New York, NY, USA.
- [23] Anna Marie Hoffman. 2021. *Parent Reports of Children's Parasocial Relationships with Conversational Agents: Trusted Voices in Children's Lives*. Georgetown University-Graduate School of Arts & Sciences, Communication, Culture & Technology.
- [24] Arvid Horned. 2020. *Conversational agents in a family context : A qualitative study with children and parents investigating their interactions and worries regarding conversational agents*. Ph. D. Dissertation.
- [25] Juan Pablo Hourcade, Natasha E Bullock-Rest, and Thomas E Hansen. 2012. Multitouch tablet applications and activities to enhance the social skills of children with autism spectrum disorders. *Pers. Ubiquit. Comput.* 16, 2 (Feb. 2012), 157–168.
- [26] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. 2004. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human-Computer Interaction* 19, 1-2 (June 2004), 61–84.
- [27] Julie A Kientz, Matthew Goodwin, Gillian R Hayes, and Gregory D Abowd. 2013. *Interactive Technologies for Autism: A Review*. Morgan & Claypool Publishers.
- [28] Phoebe Lin, Jessica Van Brummelen, Galit Lukin, Randi Williams, and Cynthia Breazeal. 2020. Zhorai: Designing a Conversational Agent for Children to Explore Machine Learning Concepts. *AAAI* 34, 09 (April 2020), 13381–13388.
- [29] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How WEIRD is CHI?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama Japan). ACM, New York, NY, USA.
- [30] Sonia Livingstone and Ellen J Helsper. 2008. Parental Mediation of Children's Internet Use. *J. Broadcast. Electron. Media* 52, 4 (Nov. 2008), 581–599.
- [31] Silvia Lovato and Anne Marie Piper. 2015. Siri, is this you?. In *Proceedings of the 14th International Conference on Interaction Design and Children* (Boston Massachusetts). ACM, New York, NY, USA.

- [32] Silvia B Lovato, Anne Marie Piper, and Ellen A Wartella. 2019. Hey Google, Do Unicorns Exist? Conversational Agents as a Path to Answers to Children's Questions. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (Boise, ID, USA) (IDC '19). Association for Computing Machinery, New York, NY, USA, 301–313.
- [33] Miriam Madsen, Rana el Kaliouby, Matthew Goodwin, and Rosalind Picard. 2008. Technology for just-in-time in-situ learning of facial affect for persons diagnosed with an autism spectrum disorder. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility* (Halifax, Nova Scotia, Canada) (Assets '08). Association for Computing Machinery, New York, NY, USA, 19–26.
- [34] Dominic W Massaro, Ying Liu, Trevor H Chen, and Charles Perfetti. 2006. A multilingual embodied conversational agent for tutoring speech and language learning. In *Ninth International Conference on Spoken Language Processing*.
- [35] Melissa Mazmanian and Simone Lanette. 2017. Okay, one more episode. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland Oregon USA). ACM, New York, NY, USA.
- [36] Simon Provoost, Ho Ming Lau, Jeroen Ruwaard, and Heleen Riper. 2017. Embodied Conversational Agents in Clinical Psychology: A Scoping Review. *J. Med. Internet Res.* 19, 5 (May 2017), e151.
- [37] Eric E Rasmussen, Autumn Shafer, Malinda J Colwell, Shawna White, Narissra Punyanunt-Carter, Rebecca L Densley, and Holly Wright. 2016. Relation between active mediation, exposure to Daniel Tiger's Neighborhood, and US preschoolers' social and emotional development. *J. Child. Media* 10, 4 (Oct. 2016), 443–461.
- [38] Dianing Ratri and Yoomi Choi. 2017. Children book app as emotional learning media. In *Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication* (Beppu, Japan) (IMCOM '17, Article 72). Association for Computing Machinery, New York, NY, USA, 1–8.
- [39] Mark Riedl, Rosa Arriaga, Fatima Boujarwah, Hwajung Hong, Jackie Isbell, and Juane Heflin. 2009. Graphical Social Scenarios: Toward Intervention and Authoring for Adolescents with High Functioning Autism. In *2009 AAAI Fall Symposium Series*.
- [40] Laurel D Riek, Tal-Chen Rabinowitch, Bhismadev Chakrabarti, and Peter Robinson. 2009. How anthropomorphism affects empathy toward robots. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction* (La Jolla, California, USA) (HRI '09). Association for Computing Machinery, New York, NY, USA, 245–246.
- [41] Hagit Sasson and Gustavo S Mesch. 2019. Parental Mediation. , 6 pages.
- [42] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I Hong. 2018. "Hey Alexa, What's Up?".
- [43] Solace Shen, Petr Slovak, and Malte F Jung. 2018. Stop. I see a conflict happening. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago IL USA). ACM, New York, NY, USA.
- [44] Kazuhiro Shidara, Hiroki Tanaka, Hiroyoshi Adachi, Daisuke Kanayama, Yukako Sakagami, Takashi Kudo, and Satoshi Nakamura. 2020. Analysis of Mood Changes and Facial Expressions during Cognitive Behavior Therapy through a Virtual Agent. In *Companion Publication of the 2020 International Conference on Multimodal Interaction* (Virtual Event, Netherlands) (ICMI '20 Companion). Association for Computing Machinery, New York, NY, USA, 477–481.
- [45] Petr Slovák and Geraldine Fitzpatrick. 2015. Teaching and developing social and emotional skills with technology. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 4 (2015), 1–34.
- [46] Petr Slovák and Geraldine Fitzpatrick. 2015. Teaching and Developing Social and Emotional Skills with Technology. *ACM Trans. Comput.-Hum. Interact.* 22, 4 (June 2015), 1–34.
- [47] Petr Slovák, Ran Gilad-Bachrach, and Geraldine Fitzpatrick. 2015. Designing Social and Emotional Skills Training: The Challenges and Opportunities for Technology Support. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 2797–2800.
- [48] Petr Slovák, Nikki Theofanopoulou, Alessia Cecchet, Peter Cottrell, Ferran Altarriba Bertran, Ella Dagan, Julian Childs, and Katherine Isbister. 2018. "I just let him cry... Designing Socio-Technical Interventions in Families to Prevent Mental Health Disorders. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–34.
- [49] Martina Smahelova, Dana Juhová, Ivo Cermak, and David Smahel. 2017. Mediation of young children's digital technology use: The parents' perspective. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 11, 3 (2017).
- [50] Abigale Stangl, Joanna Weidler-Lewis, Carlye Lauff, Emily Price, and Eric Fauble. 2017. The SEL Transition Wheel: Designing for Early Childhood Social Emotional Learning. In *Proceedings of the 2017 Conference on Interaction Design and Children* (Stanford, California, USA) (IDC '17). Association for Computing Machinery, New York, NY, USA, 334–339.
- [51] Fumihide Tanaka and Shizuko Matsuzoe. 2012. Children teach a care-receiving robot to promote their learning: field experiments in a classroom for vocabulary learning. *J. Hum.-Robot Interact.* 1, 1 (July 2012), 78–95.
- [52] Hiroki Tanaka, Hideki Negoro, Hidemi Iwasaka, and Satoshi Nakamura. 2017. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. *PLoS One* 12, 8 (Aug. 2017), e0182151.
- [53] JTanaka2016-tu Hiroki Tanaka, Sakti Sakriani, Graham Neubig, Tomoki Toda, Hideki Negoro, Hidemi Iwasaka, and Satoshi Nakamura. [n. d.]. Teaching Social Communication Skills Through Human-Agent Interaction. ([n. d.]).
- [54] Julia S Y Tang, Marita Falkmer, Nigel T M Chen, Sven Blte, and Sonya Girdler. 2021. Development and Feasibility of MindChip™: A Social Emotional Telehealth Intervention for Autistic Adults. *J. Autism Dev. Disord.* 51, 4 (April 2021), 1107–1130.

- [55] Monica Tentori and Gillian R Hayes. 2010. Designing for interaction immediacy to enhance social skills of children with autism. In *Proceedings of the 12th ACM international conference on Ubiquitous computing* (Copenhagen, Denmark) (*UbiComp '10*). Association for Computing Machinery, New York, NY, USA, 51–60.
- [56] Sarah J Tracy. 2010. Qualitative Quality: Eight “Big-Tent” Criteria for Excellent Qualitative Research. *Qual. Inq.* 16, 10 (Dec. 2010), 837–851.
- [57] Wayne Ward, Ron Cole, Daniel Bolaños, Cindy Buchenroth-Martin, Edward Svirsky, and Tim Weston. 2013. My science tutor: A conversational multimedia virtual tutor. , 1115–1125 pages.
- [58] Iain Werry, Kerstin Dautenhahn, Bernard Ogden, and William Harwin. 2001. Can Social Interaction Skills Be Taught by a Social Agent? The Role of a Robotic Mediator in Autism Therapy. In *Cognitive Technology: Instruments of Mind*. Springer Berlin Heidelberg, 57–74.
- [59] Ying Xu and Mark Warschauer. 2019. Young Children’s Reading and Learning with Conversational Agents. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (*CHI EA '19, Paper CS10*). Association for Computing Machinery, New York, NY, USA, 1–8.
- [60] Joseph E Zins. 2004. *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* Teachers College Press.

## A APPENDIX

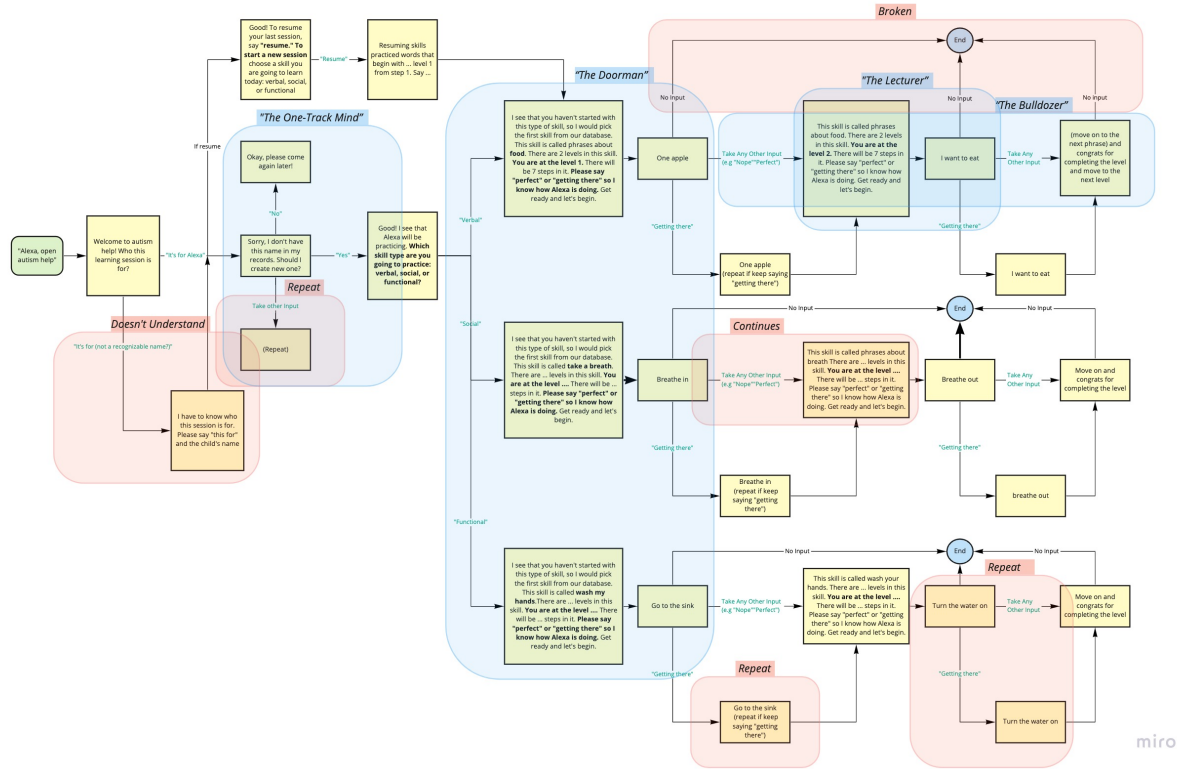


Fig. 5. A interaction diagram of the Skill “Autism Help”. Blueshading highlights interaction styles, and red shading highlights recovery strategies.