Evaluating an iPad Game to Address Overselectivity in Preliterate AAC Users with Minimal Verbal Behavior

LouAnne E. Boyd[1], Kathryn E. Ringland[1], Heather Faucett[1], Alexis Hiniker[2], Kimberley Klein[1], Kanika Patel[1], Gillian R. Hayes[1]

Department of Informatics
University of California, Irvine
{boydl, kringlan, hfaucett, knklein, kanikap, gillianrh}@uci.edu

Information School
University of Washington
alexisr@uw.edu

ABSTRACT

Overselectivity is a learning challenge that is largely unaddressed in the assistive technology community. Screening and intervention, done by specialists, is time-intensive and requires substantial training. Little to no treatments are available to the broader population of preliterate, minimally verbal individuals. In this work, we examine the impact of an iPad game based on the tenets of behavioral therapy to mitigate overselectivity. We developed software-based techniques and evaluated the system using established methods from the field of Special Education. We present the results of a deployment in a special education school that demonstrates that an assistive tablet game is a feasible means of addressing overselectivity, and we present generalizable technological features drawn from evidenced-based therapies to consider in future assistive technologies. We suggest that designers of assistive technology systems, particularly those who address physical, cognitive, and behavioral difficulties for preliterate AAC users, should consider overselectivity as a potential co-occurring condition.

CCS Concepts

• Human-centered computing → Accessibility → Accessibility technologies

Keywords

Autism; AAC; assistive technology; children; tablet games; language development; multiple cue responding; overselectivity.

1. INTRODUCTION

Despite rapid advances in technology, the ability for those who depend on Augmentative and Alternative Communication (AAC) devices to express their wants and needs through such technologies is still highly limited due to a variety of challenges. These devices require intensive setup and often support only the most basic needs [1], such as requesting items or activities. Additionally, commercially available devices, although somewhat customizable, still fail to meet the highly-individualized needs of many AAC users, creating an opportunity for creative solutions.

Improving interactions with these systems has been a continual area of interest for both researchers and practitioners, and assistive technology specialists have identified several potential challenges for AAC users. For example, those with cognitive difficulties may not form sentences as expected [30]. Similarly, those with memory problems (such as aphasia) may not remember how to use the system [1]. AAC devices are restricted to the corpus and categorization provided by the designer, and designs may be limited for those with certain physical disabilities [31]. Finally, because symbols mean different things to different people, these systems do not necessarily work across cultures and contexts. Several studies have focused on making prediction faster because speed of communicating through the device may be prohibitive and lastly, literacy can be an issue for systems that require it [29]. Furthermore, such challenges are often disability specific, leading to more specialized designs. While such specialization is important, this focus does not address challenges that are more general yet still pervasive.

Individuals with pervasive developmental delay often struggle to see the proverbial forest for the trees, meaning that they miss the gestalt overview in exchange for a focus on details, known as "overselectivity" [20,24]. This common learning challenge for students in special education, often called simply "overselectivity," results in overly narrow attention to salient stimuli in the environment and impacts one's ability to learn from observing the environment. Until recently, this phenomenon had been primarily described in children with autism; now, there is growing evidence overselectivity impacts a broad range of people such as people with developmental disabilities, learning disabilities, and hearing impairments [4], many of whom may be AAC users. Overselectivity has also been observed in preliterate AAC users who may not be identified for treatment [5].

Currently, there are very few methods for identifying and treating overselectivity, and the methods that do exist are expensive, time-intensive, highly specialized, and only available for specific populations [11,12]. The pervasive and highly impactful nature of overselectivity provides a compelling reason to develop feasible and accessible treatments to reduce overselectivity in AAC use. The possibility that a larger percent of preliterate AAC users could display overselectivity provides a compelling reason to design systems to reduce overselectivity. Improving issues with overselectivity should improve long-term accessibility as AAC systems are upgraded, often resulting in changes to appearance and to the function of interfaces [5]. We present results from a deployment study of a suite of games, *Go Go Games*, designed by the fourth author to reduce overselectivity in children through behavioral therapy [7]. The results of our evaluation of *Go Go Games* at a special education school demonstrate the potential of reproducible techniques to reduce overselectivity for preliterate, minimally verbal children who display poor uptake of AAC systems.

ASSETS'17, October 29–November 1, 2017, Baltimore, MD, USA.

DOI: http://dx.doi.org/

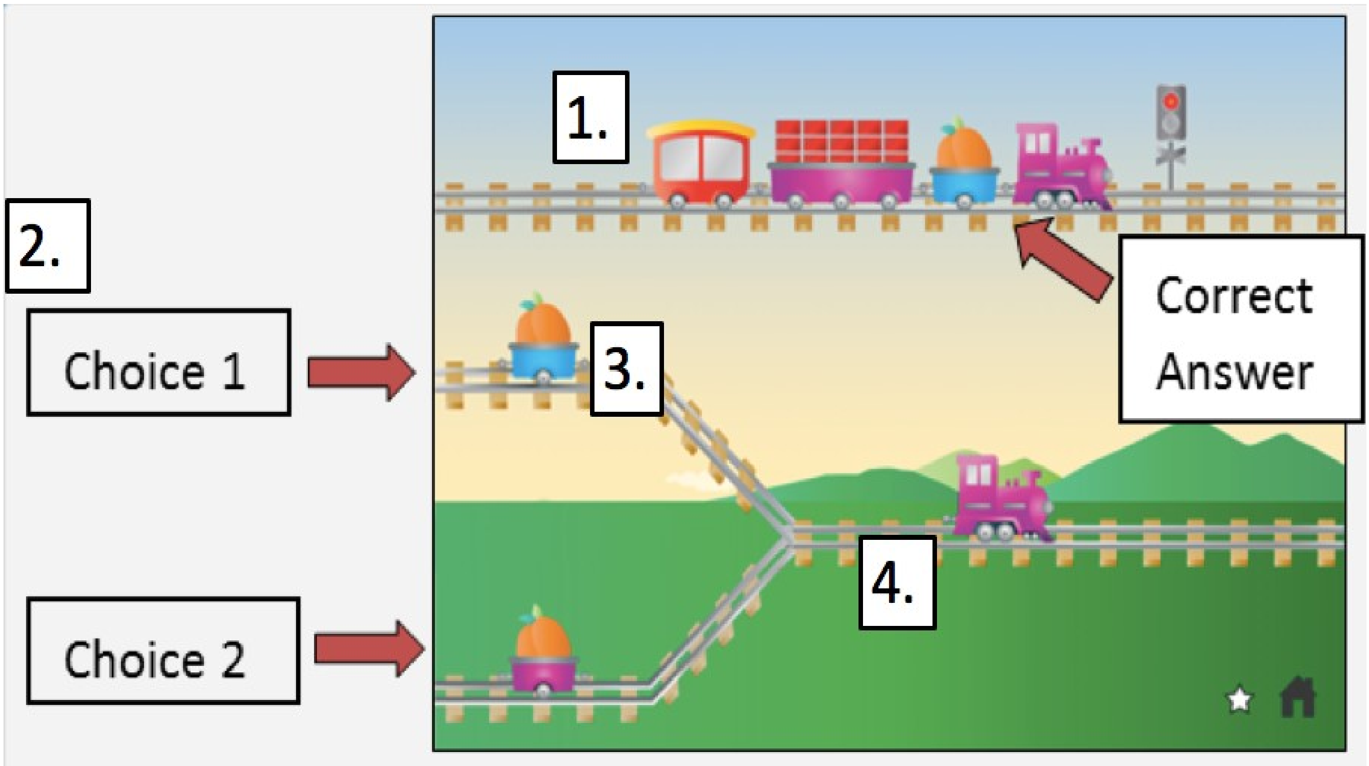

Figure 1. Steps 1-4 show the user interaction in the train minigame: 1. Observe graphic prompt,2. Scan multiple options, 3. Select matching option, and 4. Drag and drop selected choice. © Go Go Games Studios, LLC

2. RELATED WORK

The opportunities to improve the usability and accessibility of AAC devices revolve around a wide variety of innovative ways to adapt the interaction or information (i.e., language set) to support a user's expressive communication. Yet there is an underserved need for assistive technologists to implement replicable techniques to address overselectivity. These techniques, if shown to be feasible and effective, could then be implemented across assistive technology devices to support a range of conditions. To understand the innovative work that has been conducted regarding AAC systems for preliterate, minimally verbal communicators in the ASSETS community, we have found the focus to be predominantly on simplifying the interaction.

Many of these systems incorporate the use of icons as an alternative to verbal, gestural, or written communication. Paper-based systems have relied on icons to support expressive communication for decades [22], and with the affordability and portability of mobile devices, icon-based systems have expanded [15]. Several researchers have extended the utility of icon-based systems such as adding multimedia to the system (e.g., photos, video). For example, Prior et al. [23] developed the *CHAMPION* software project that supported minimally verbal communicators to expressive themselves using–multimedia in the contexts of hospitalization so where unfamiliar staff could understand the user [25]. Similarly, *Vid2Speech* supported early communicators who are preliterate and preverbal by adding video of action words to represent their "dynamic and ephemeral" nature [18]. These systems extend the type of interactions users can have with systems.

Another approach to supporting users with communication is to provide an intervention rather than an alternative form of communication. For example, Black focused on preliteracy skills by developing *PhonicStick*, a joystick that blends and outputs letter sounds to allow for minimally verbal children to explore phonics in an effort to advance emerging literacy skills [2]. Our work also addresses a learning challenge frequently found in preliterate, minimally verbal people--overselectivity.

2.1 OVERSELCTIVITY

Overselectivity to non-relevant or isolated aspects of the environment interferes with children's ability to learn to use language flexibly. Specifically, overselectivity interferes with one's ability to attend to discriminations in language, a skill children typically acquire around the age of three or four. For example, by this age, children can typically respond correctly to the request, "show me the green ball," which requires attention to the object (ball) and color (green). Dube and Wilkinson define stimulus overselectivity as

> *"an atypical limitation in the number of stimuli or stimulus features within an image that are attended to and subsequently learned…For example, the Mayer Johnson PCS symbol for TENNIS shows a gray racquet with a yellow ball…if overselective stimulus control were restricted to the ball only, and the student had learned to identify the symbol on the basis of that one isolated feature alone, then the student may make errors during subsequent symbol use when the symbol BALLOONS is present because that symbol includes a yellow balloon about the same size and color as the tennis ball" [5:4].*

Overselectivity impacts learning in social, speech, and observational learning, as well as intellectual development [26]. Overcoming this barrier is believed to create a path to children accessing more inclusive learning environments [27]. Therefore, determining the therapeutic potential of this technology is critical to understanding a new role technology can play in supporting access to everyday life through improved use of AAC devices and comprehension of language. This project offers both insight into the potential for technical assistance to bridge a gap in accessing existing communication devices, highlighting technical features that can be parlayed into a variety of devices. Support is then sustained over the inevitable changes users experience in devices as interfaces get upgraded or users move on to other systems. This work aims to identify technical features to teach multiple cue responding and mitigate overselectivity. Overselectivity can be problematic for users who move across assistive technology devices, therefore reproducible technical features to mitigate overselectivity, are important.

3. MULTIPLE CUE RESPONDING

Multiple Cue Responding (MCR) is the ability to observe and attend to multiple cues (e.g., color and shape) and recall those cues to make decisions (e.g., find the red triangle in a field of a red square, a blue triangle, and a red triangle). Teaching MCR helps children learn to communicate verbally and may help children use their AAC devices [5]. Pivotal Response Treatment (PRT), one of the only current treatments that teaches MCR, is a therapeutic approach focusing on a child-initiated perspective to target “pivotal” skills that are known to bring about improvements in social, communication, and academic domains [12]. This approach blends the goals of caregivers with the interests of the child to create opportunities for the child to express his or her wants and needs. PRT harnesses the child’s motivation to communicate and systematically introduces more complex environments that require the child to then use multiple cues to respond effectively [11].

To examine how the key principles of PRT translate into an iPad game, harnessing children’s motivation and systematically exposure children to multiple cues, we developed *Go Go Games* [7]. In this work, we deploy *Go Go Games* to determine the effectiveness of these techniques in assistive technologies which is critical to understanding the role technology can play in MCR intervention. Teaching MCR helps people learn to communicate verbally and also may help children use their AAC devices [5].

3.1 *Go Go Games*

Go Go Games is a publicly available suite of therapeutic video games designed around the principles of Pivotal Response Training (PRT) [12,16]. Players use the system to work through PRT-grounded exercises at their own pace. Players engage with child-preferred digital objects to build trains, direct cars, and assemble robots on screen. To find the correct pieces, they must pay attention to multiple features of the items they see, such as the length, color, and cargo of the train cars they choose. The user interacts with the tablet game by first observing the stimuli presented (i.e., the graphic prompt); 2. scanning options, 3. selecting a match, and 4. dragging selection along path to complete the trial (see Figure 1). Generalizable techniques translated into digital form include:

Repetition of brief user interaction: In *Go Go Games*, users observe, scan, select, drag and drop an item to the correct location along a visual pathway (e.g., train track).

Feedback & Progress Monitoring: In face-to-face therapy, physical token boards represent how much work is expected; this token analogy is visually replicated in the *Go Go Games* interface (see star icon in bottom right corner of Figure 1). Corrective feedback was given by blinking the perimeter of the correct image after the input of an incorrect answers, a progress tracker on the initial screen and in the corner of each game screen, and a congratulatory screen when each level was completed.

Activity is Child-led: *Go Go Games* incorporates child-preferred stimuli (e.g., iPads, trucks, cars, spaceships). Users have flexibility to choose which game and level to play and can turn on or off background music.

Systematic Intervention: The software systematically scaffolds task difficulty based on user performance. Each trial randomizes the placement of correct options on the screen. Immediate feedback is provided for each trial (i.e., visual and audio statements are differentiated based on response).

An additional aspect of a therapeutic iPad game is, unlike traditional therapies, no other person is required to be part of the interactions. We hypothesized that the unique features of *Go Go Games* would support learning MCR as well as increase child motivation to participate in teaching MCR.

4. METHODS

To evaluate if *Go Go Games* can support MCR, we conducted an experiment in a special education school. We conducted this work with school staff providing 6 students with up to 10 minutes of play per day, across nine days (group avg.= 66 minutes, range 30-90 minutes). To ensure that only the game was available during testing sessions, staff reported they enabled a technical feature of the iPad, “guided access.” Staff logs were used to calculate minutes of game play, and comments were reviewed for themes about the participants’ engagement with the game. Concurrently, we conducted assessments of MCR in the physical world. After the study, we collected usage logs recorded by the teachers and conducted interviews of the teaching staff (n=6).

4.1 Study Design

We employed a single-case experiment design in a special education classroom, as is widely used by special education researchers and clinicians [17]. This evaluation technique is important to the special education community as a body of evidence about the effectiveness of assistive technology in context of use is critical to the inclusion of therapeutic tools in the classroom. Single-case research uses a single participant as one’s own control and aims to replicate desired effects across 3-8 participants per study [9]. Additionally, the common practice is to present data on “responders” as well as “non-responders” to add to the knowledge about a specific population [9].

Two of three classroom teachers volunteered to participate in the study, and all students in those two classrooms were invited to participate. After a parent orientation meeting provided by the first author, eleven students’ parents gave consent for their children to be in the study. Next, we randomly assigned one class to play the game first (5 students), while the other class played second (6 students). Assessors were blind to which group the participants were in and the amount of game play they received during the study. We employed visual analysis to interpret the assessment data on a case-by-case basis [6,25].

Each participant began the study with a three-day baseline assessment of their MCR skills in the physical world. Group A then began nine days of *Go Go Games* play, while Group B maintained regular classroom activities. After this phase, both groups were measured again, resulting in a post-intervention assessment for Group A, and a repeated baseline measure for Group B. Then Group B began 9 days of *Go Go Games* play while Group A resumed regular school activities. Both groups were measured at the end for a post measure for Group B and a follow up measure for Group A. The strength of this design in an applied setting is that the change that would result from maturation alone is controlled for by staggering the conditions over time [3]. Therefore, if changes are observed between phases, and across multiple baselines, the likelihood the change is due to the intervention is increased. Additionally, if improvements are maintained after the intervention has been removed during the follow up assessment, these changes can be attributed to the intervention.

The staff designated a ten-minute period per day to play *Go Go Games* and logged the start and end time for each student as well as any comments about the interaction. Participants were assessed during the regular school day at the beginning of the study, at the mid-point (between two groups), and at the end (see Table 1).

We assessed the participants' MCR with physical world objects during three phases of the study (beginning, middle, and end) replicating procedures described in behavioral research literature [24]. Assessments occurred one participant at a time in a separate room. Assessment sessions lasted from 3-30 minutes depending on the behavior of the child. We met with the school's behavior therapist for the first 3 days to align our reinforcement schedules to those being used in the classrooms to ensure student comfort. We also aligned our protocol for responding to undesired behavior (e.g., screaming, spitting, throwing objects, hitting, falling to the floor) to that of the school to minimize disruption to their intensive behavioral programming in place for each student.

This resulted in continuing the assessment while engaging in the planned ignoring of disruptive behaviors (i.e., screaming, spitting*),* and terminating the session for aggressive behaviors (i.e., hitting or non-responsiveness to instructor for 3 minutes).

Table 1: Study design

Days/ Group	1-3	4-12	13-16	17-25	26-28
A	Baseline	Gameplay	Post test	Wait	Follow up
B	Baseline	Wait	Repeated Baseline	Game Play	Post test

4.2 Participants

The school field site provides highly individualized education and therapy during the school day to support academics, functional life skills, social-emotional, and physical development.

One behavior therapist, two lead special education teachers, and six teaching assistants participated in the study as instructional staff. The parents of eleven out of fifteen students with multiple learning challenges, age 6–14, consented to participate. Eleven participants began the study by participating in the baseline screening assessment of MCR (see Table 2). Not all participants were eligible as three demonstrated perfect scores on the screener (P4, P9, P10) and the school administrators had to remove research activities from the schedule under a variety of contexts (i.e., P6 exhibiting aggressive behavior or was not available due to a therapy session). Additionally, measures were missed on multiple days when specific children were absent from school (P7, P8). We present the assessment results of 6, all AAC users, some of whom missed some assessment sessions or game play.

4.3 MCR Probes

To understand the impact of playing *Go Go Games*, we compared performance on a physical MCR assessment before and after the children played the game. The physical assessment and training for these skills incorporate common classroom objects to test the ability to discriminate among objects based on features in their environment [24]. The assessment tasks consisted of fourteen levels of matching 3D objects that grew in complexity from matching dissimilar wooden objects such as a red square and yellow triangle to recalling textures of a red rectangular block, and finally to distinguishing which image was presented between similar 2D images. The final level used physical flashcards we made from digital images in *Go Go Games* as an additional prerequisite level of testing.

The researchers, board certified behavior analysts, and trained assistants, conducted the assessments based on MCR assessment procedure. A researcher holds up a block in front of the child and states, "this one is the correct one" then removes the object from view. A moment later, the researcher places the correct item and its pair in front of the child, holds out her hand for the participant to place the correct block there and then states, "give me the correct one." Before the test starts, the researcher teaches this response by prompting (e.g., pointing at the correct answer) and then moves on to the test phase. In the subsequent trials, the child is expected to independently makes the distinction of the correct block based on differences in features (i.e., shape, color, size, texture, shade, and finally multiple features in 2D images). The school staff requested an additional level of testing be introduced that permitted the items to remain present during the trial. These two phases resulted in 14 levels.

We worked with the school staff to collaboratively determine if the child needed an additional reinforcement system and to receive advice about engaging the child. For example, P1 required an extended interval for her response time given her motor challenges and P5 had a known fixation with items that are red, thus alternative colored blocks were used. During the assessments, two researchers collected data on correct or incorrect responding, and they tallied the scores daily. The complete administration of the test ranged from 5-15 minutes each time and took place in a separate designated room in the school.

4.4 Data Analysis

To measure effectiveness of the game as an intervention for teaching MCR, we compared the distance for each trial between baseline and treatment, a procedure known as "Nonoverlap of All Pairs" (NAP) [14]. "NAP is a 'complete' nonoverlap index as it individually calculated as the number of improving or positive (Pos) pairs plus half of ties (.5 × Ties), divided by all pairs: NAP = ([Pos + .5 × Ties] / Pairs)" [26]. This nonparametric measure of treatment effectiveness is a common method in behavioral research using single-case experiments for autism interventions. This approach is necessary given the small sample sizes typically present in autism intervention research. Using the guidelines for interpretation recommended by Parker and Vannest, NAP scores between 0 and .65 can be classified as "weak effects" (i.e., no effect), 66 to .92 as "medium effects," and .92 to 1.0 as "strong effects" [19].

Table 1 : Participant demographics and MCR assessment results

Group	ID	Age	Mode of Comm.	Edu. Label	Gender	Baseline	#of play sessions	Avg. Mins of usage	Total Mins of usage	%NAP	Effect Size
A	P1	12.8	AAC	ID	F	4/14	9	10	90	22	Weak
A	P2	11.1	AAC	ID	M	6/14	9	10	90	83	Medium
A	P3	6.2	Verbal	AUT	M	14/14	screened out				
A	P4	6.9	Verbal	AUT	M	14/14	screened out				
A	P5	7.3	AAC	AUT	M	0/14	8	8	65	100	Strong
B	P6	13	Verbal	AUT	M	2/14	study terminated for aggression				
B	P7	11.1	AAC	AUT	M	9/14	6	10	60	33	Weak
B	P8	10.7	AAC	ID	F	0/14	3	10	30	54	Weak
B	P9	8.11	Verbal	AUT	M	14/14	screened out				
B	P10	11.5	Verbal	AUT	M	14/14	screened out				
B	P11	11.6	AAC	ID	M	3/14	8	8	66	88	Medium

To understand the experience of using the game, we analyzed data from the six staff interviews alongside our observational notes. The research team individually reviewed the interview notes for initial impressions and then met collectively to discuss trends in the data. Collectively the team then sorted comments into topics. After identifying dominant topics, we re-sorted the comments into themes. We compared these themes with the effectiveness results and presented these findings to the school's behavior analyst for her interpretation of the collective data. With her input, we then analyzed our observational data for evidence for or against these themes in an iterative manner, then we aligned themes with interactions with the software features.

5. RESULTS

We present results in Section 5.1 that indicate that *Go Go Games* was at least moderately effective in improving some students' ability to consider more stimuli in a task. We analyzed interview and observational data from all participants, resulting in findings related to the re-usable interactive features of *Go Go Games,* presented in Section 5.2.

5.1 MCR Assessment Results

Of the eleven participants who were screened, nine qualified and began the study, and six completed enough trials for analysis: three from group A (P1, P2, P5), and three from the extended baseline condition, group B (P7, P8, P11). we present a description of each of the six participants who participated in each phase of the study (excluded from the study were: P3, P4, P9 and P10 who demonstrated mastery of MCR in the screener and P6 who was dismissed from assessment for aggression toward the staff). For these six cases, we conducted a visual analysis [6,25] of the time series data and treatment effect with the non-overlapping procedure described above (see Table 1). A vignette of each participant case assessment is described below.

P1, a 12-year-old girl, has a combination of autistic-like characteristics and a genetic condition called Phelan-McDermid syndrome which often results in delays or impairment with cognition and motor skills [19]. She is minimally verbal and carried an iPad for expressive communication. Her AAC device contained a few choices of words and phrases that she could press on when she wanted to express a feeling or speak about a certain object, such as requesting a snack. Staff described her *Go Go Games* play as prompted due to her delayed motoric responses, and so we permitted extra time during the MCR assessment. P1 played *Go Go Games* over 9 days for 10 minutes. Her MCR assessment results revealed a low level, yet slightly increasing trend in baseline. After nine days (90 minutes of intervention), she exhibited a lower level and slightly more variable trend that persisted during her follow up assessment (see Figure 2). P1's case shows a weak treatment effect as the data only displays a 22% nonoverlap (see Table 2). Because there is no difference between baseline and post-intervention assessments, we conclude the game did not impact her ability to respond to multiple cues in her physical word.

P2 is dually diagnosed with autism and Down syndrome. Staff explained that when he played, they had to prompt him to continue. Visual inspection of P2's baseline assessment shows a low level of MCR with an increasing trend, resulting in a moderate level of performance. P2 played *Go Go Games* over 9 days for 10 minutes. In contrast, P2's post-intervention assessment showed more variability with higher level of performance. During his follow-up (after a nine-day break from his game play), he maintained a moderate level of performance with some variability that is similar to his baseline but at a slightly higher level. Given his two high scores post intervention, his NAP score was 82%, indicating a medium treatment effect as some of the post treatment scores exceed baseline scores.

P5 is a minimally verbal boy with autism and intellectual disability with a reported history of aggressive behavior. Upon entering the assessment room, he laid on the floor. He grabbed a red block and said "R" for the color red. We were told he perseverates over objects that are red so we switched to yellow blocks. On occasion, he signed "all done" and we ended the trial. At times, he would hold the blocks over his head in a gesture that looked like he would throw the blocks. He matched correctly until the second half of the assessment where the task shifts from matching with the cue present to matching from recalling the color of the block. Staff reported that he played the game independently with frequent requests to continue playing at the 10-minute timer. He fluidly moved between mini-games at times taking a break to engage in self-stimulation. Through visual inspection of P5's performance, he scored zero correct in his three baseline sessions. After the intervention phase, his scores varied at moderate levels and were replicated in follow up probes (see Figure 2). P5 exhibited strong positive treatment effects. In baseline, he did not complete even the first level designed to introduce the task. In the post intervention phase, he scored in the mid-range but with less stability and in the follow up sessions, he maintained these mid-level performances. This outcome results in 100% of NAP, a strong effect as all the post treatment points exceed the baseline points (see Table 2).

P7 is a minimally verbal, eleven-year-old boy with autism who demonstrated mastery of the first half of the MCR assessment on the baseline screening. Staff commented on his brief periods of game play "(he) played for 2 minutes then tried to exit the game, teacher redirected.-A staff member from Group A said in the post study interview that, "*he liked it, he had trouble when he hit a harder level go back to an easier level, something he was comfortable with. He would perseverate on the button choice. He would hit me he doesn't like being wrong. When it highlighted another choice, I would hand-over-hand to the correct choice and he didn't like that*." According to the staff log, he played for 10 minutes per day for the first 6 days. He was absent form school for 1-2 days per week during the duration of the study. His post intervention assessment scores are like his baseline scores except for a dip on the last day of assessment from level 7 to 2, resulting in a weak treatment effect, and a 33% NAP. Our fieldnotes reveal he performed the first two levels without error, then he began screaming, grabbed himself, and ran to the bathroom.

P8 is minimally verbal girl with intellectual disabilities. At the onset of the study she had only attended this school for two weeks and was displaying aggressive behavior. The two baseline sessions she was present for were terminated within 3 minutes for noncompliance. In the extended baseline, she participated until she began to err at level 2. Our fieldnotes indicate she continually chose the block on the left side. In the post sessions, she refused to let go of the block or threw it resulting in 2 of 3 sessions being terminated in first minute resulting in a weak treatment effect, and NAP of 54%. Her log entries for gameplay reveal she played a total of 30 minutes (10 minutes a day) on days 2,5,7 of the nine-day stage. Staff commented in the log that she was distracted, tried to exit out of game, was a bit frustrated and had some difficulty attending. A staff member from Group A reports "*she would not sit for a long, 30 seconds. She concluded stating she "isn't too reinforced by the iPad*."

P11 is also a minimally verbal boy with autism and cerebral palsy. He is very interested in socializing with staff as he moves throughout the school and stops to visit and chat with others through his iPad. In assessment sessions, he pushed buttons that output "sing happy birthday" & "Where is [staff name]." He was preoccupied with trying to turn on the TV. Staff reported he was usually distracted by the presence others in the room and opportunities to socialize through approaching and greeting. Upon visual inspection on P11's performance, we see his baseline phase was repeated, as he was part of Group B. In the first baseline series, his scores were of a low level and the trend was flat, suggesting the trend was stable. In the second baseline, there was some variability as the middle session shows a mid-level score with no clear trend. These mid-level scores are replicated in the post-intervention assessment and are more stable. His NAP is 88%, a medium treatment effect (see Table 2).

Taken together, the analysis across these six participants indicates that a small dose of intervention, through this tablet-based game, is associated with measurable improvements in MCR performance for minimally verbal children in our sample who use AAC devices. Specifically, three of the six children who received the therapy for the full experimental dosage achieved medium or strong treatment effects (P2, P5, P11).

5.2 Implementation Results

For the participants who did respond positively to playing *Go Go Games,* we explored the generalizable techniques: preferred stimuli (the form factor of an iPad and game content); the simple interaction, and the systematic feedback. Regarding the form factor, our findings reveal variation in the children's inherent responses to the iPad game and the extent to which it could serve as a fun, reinforcing activity. Staff members that worked directly with the respective participants reported that while some "*really enjoyed it and always wanted to keep going after their time was up*," (P5, P9, P10), other participants were "*not reinforced by the iPad*" (P6, P7, P8). Only one participant verbally responded to the interview question, what did you think of *Go Go Games?* and he, (P10), said "I like trains." These features appeared generally preferred.

Regarding the simple, repetitive user interaction, responses varied between responders and non-responders. The interaction permitted immediate independence for some of the participants. Staff reported that P5 learned interaction easily – "after three prompts P5 understood the game and followed the instructions exactly". Classroom staff reported that participants (P3, P4, P11) "really engaged with the game. The game is exciting to play, [they] engaged with moving the car." Another staff member from group B reported "it's really good because of the repetition." As the iPad game had appeal to the responders, the game mechanics ultimately impacted the success of a student.

One of the biggest indicators of success for this group of participants was the child's tolerance for feedback – both in game and from staff. Staff reported some participants, such as P7, "didn't like being wrong." Therefore, the corrective feedback became a critical feature. Frequently, the staff logs indicated that students were prompted by the staff to continue to play for the suggested 10 minutes, and, in the case of P1, physically assisted to drag the icon to the correct spot once she touched an object. P11 also received assistance to play the game with a staff person sitting with him at times to encourage him to stay seated and playing the game. Children who were willing to receive help from the staff were more likely to complete the requested amount of playtime and repeated assessments.

As with receiving feedback, difficulty stopping play can be a problem for many children. Difficulty transitioning between activities can be very challenging for children with autism [10], sometimes resulting in aggressive or destructive behaviors (i.e., hitting others, throwing a device). Staff reported aggression from P5 when he was asked to hand over the iPad. There is a burgeoning interest among designers to reduce tantrums from turning off tablets by leveraging natural endings [8]. Research states that some caregivers report that they withhold access to technology to avoid battling when it is time to put it away providing natural stopping points and a definitive conclusion has the potential to make the experience more accessible to more families [15]. Addressing this challenge with transitioning away from using the iPad would be a next step to help students like P5 to continue to learn MCR.

Ways we suggest that therapeutic games for this population could be programmed to come to an end smoothly (e.g., the screen could begin fading until it is blank) to make the transition to disengage

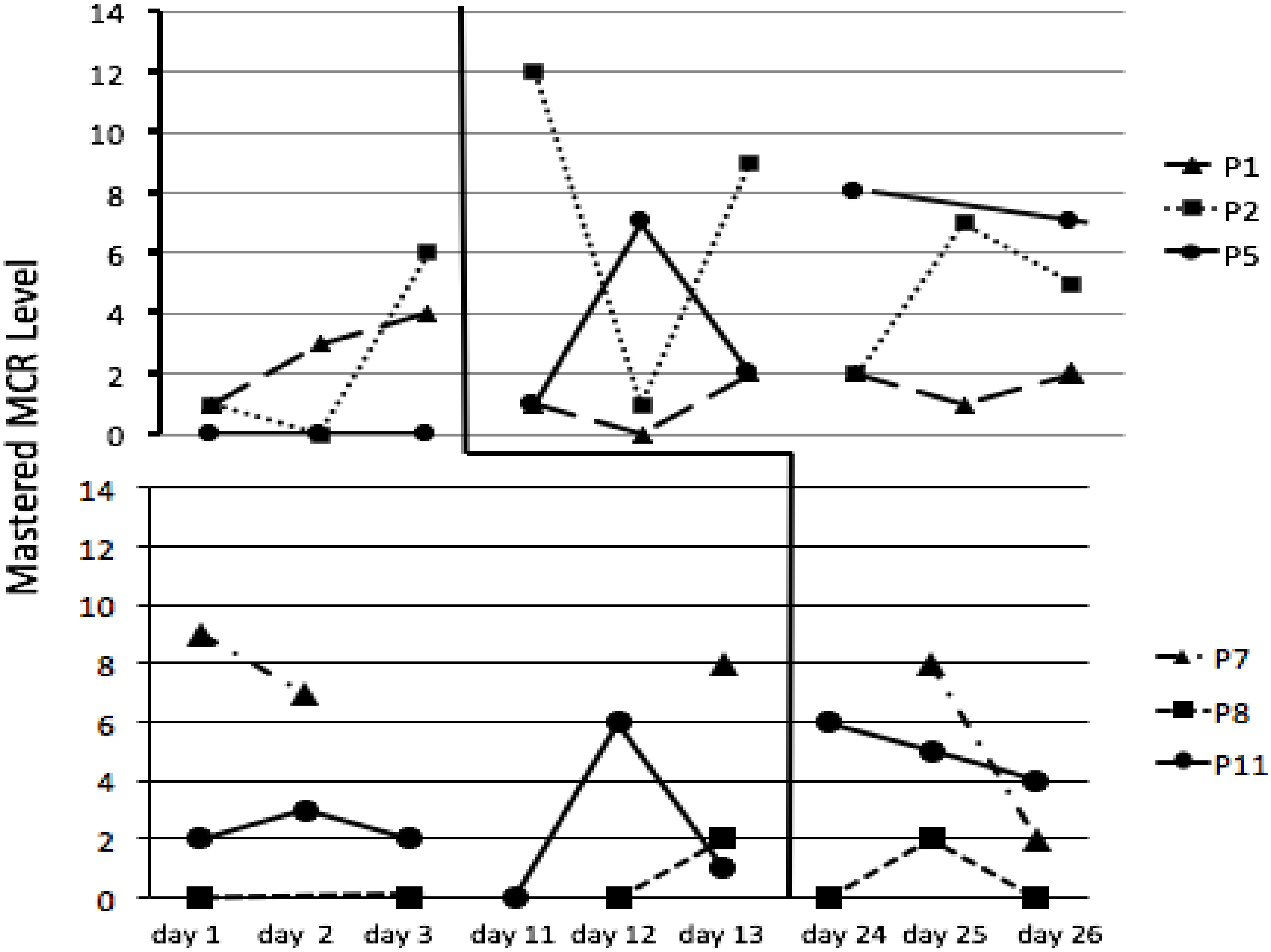

Figure 2: Participant's level of MCR before and after intervention.

more natural. The guided access application used at this school removed the option to exit the game and enter another application but did not address the transition to ending gameplay. An automated way to end play could alleviate some of the discomfort in transitioning (i.e., created a natural ending [8]). Applications can be designed to specifically support terminated iPad game play, methods to do so that do not lead to aggression would require future work.

5.3 Managing Co-Occurring Conditions

Given the complex and varied needs of the AAC users in our study that required or avoided human prompting or sensory input, we present three design considerations to support varied presentations for children who engage in overselectivity, specifically as they related to physical, cognitive, and behavioral challenges. Our analysis highlighted the complexity of the context of classroom settings, in which we deployed *Go Go Games* and which are likely targets for therapeutic game use more generally.

5.3.1 Physical Disabilities

Physical disabilities that co-occur with overselectivity can cause additional challenges, both related to overselectivity and related to the intervention used to improve it. Several staff mentioned visual concerns, for example, saying that the games presented too many details or that the digital objects were too small or too subtle. or making specific suggestions:

"Making the flags or pictures bigger may be helpful for students to see." - Group B staff member

Similarly, our observations and staff reporting indicated that motor issues were a concern for P1 who has a motor disability, who could not easily use the game despite it being designed to tolerate imprecise touch input, allow breaks in smooth drags, and support players who struggle to perform motions that require crossing the midline of their bodies.

These results indicate that all games, and *Go Go Games* should be further modified to allow for additional types of interactions. For example, a touch or tap, instead of a drag would simplify access for children with motor issues. Similarly, personalization of the visual output, as suggested above, could improve access for those with visual challenges. Although wide ranges of input have been explored previously in the literature [28], what is key to our findings here is the challenge of intersectionality within the assistive technology and disabilities communities. It is not enough to design systems with one or even multiple common combinations of disabilities in mind. Children are often evolving in their disabilities experiencing dynamically changing multiple conditions and contexts, requiring intense personalization while still, ideally, making the system accessible for use without professional configuration. This requirement speaks to the need for future therapeutic games to employ intelligence to learn a user's skillset and capabilities as well as to employ personalization at a level that a child, parent, or teacher could engage [32].

5.3.2 Cognitive and Attention Related Disabilities

Just as co-occurring physical disabilities can greatly increase the complexity of a child's struggle with overselectivity, many children in our study also experienced additional *cognitive* and *attention* issues that may have played a role in their engagement with the game. These issues are challenging to differentiate from overselectivity but worth exploring.

As just one example of these challenges, for people to use *Go Go Games* successfully, they must be able to complete a swipe motion. Which can be a challenging concept to learn while playing the game. For P1, staff expressed concerns that this interaction was too cognitively demanding for her but suggested a simpler interaction might help as she does tap her AAC device. Staff also explained that several participants struggled with attention challenges when playing. A staff member from Group A explained that some children "*engage for the whole time, while it was hard for others to maintain engagement.*" Given this variability, it is not clear that all children would be able to attend to the game long enough to see therapeutic benefits.

The context of the environment of assistive technology use can also greatly impact the ability of children to engage. The work in this study was all completed in the classroom environment, in which multiple children of mixed abilities were often present. Staff reported that distractions in the classroom contributed to poor attention and that small visual details in the game were too subtle to hold children's attention. They also explained that some participants struggled to direct their attention to the focal point within the game, and for example, got stuck on one action or spot, or had difficulty understanding where the action was to occur. Others would choose to replay a successful level rather than move on to the increased challenge of the next level. A staff member from Group B suggested that, "*maybe a simpler beginning stage in the game then fade up to full screen of images,*" would provide better scaffolding for learning MCR and support children with attention challenges.

Separately, staff reported that for some participants the background images embedded in the game were distracting. Some staff suggested using minimal visuals on-screen, so that the game would be more accessible to students who are overwhelmed easily. Relatedly, some children engaged with the screen in unintended ways. Since P5 could use the iPad independently, he was observed to engage in self-stimulation with the game where his focus did not appear to be on practicing his matching skills. Despite these unintended interactions, his MCR performance improved dramatically in the post assessment. Perhaps having the opportunity to engage with the game independently, as he chose, afforded him the time to self-regulate. Debates about the appropriate role of educational technology continue in special education settings [13]. Finding a balance between engagement and distraction remains an ongoing challenge and the most productive balance may differ from one child to the next.

These results indicate that there are multiple ways children engage with technology and each way should be considered carefully to determine what, if any, modifications to a system need to occur to support accessibility.

6. DISCUSSION

The results of this research demonstrate that *Go Go Games* can be both usable and effective for learning multiple cue responding for some children with minimal verbal abilities and intellectual disabilities. Our empirical study in an applied setting provides understanding of user experiences across a small yet diverse set of users, leading us to broader questions about the nature of designing assistive technology. In this work, we demonstrated that *Go Go Games* was usable for some in this context. Our small group of diverse participants engaged in scheduled, dose-regulated game play, a necessary precursor to successful therapy. School staff could administer the game doses successfully, another essential component of creating an effective experience. Our interviews with staff members and observations of participants suggest that the app can be deployed as a therapeutic tool in a school setting. However, the fact that several of our participants dropped out before receiving the full treatment also suggests that *Go Go Games* may not be a feasible treatment option in all contexts or for all children given their individual physical, cognitive, and behavioral needs. However, given the fact that the systematic teaching of MCR currently only reaches a very limited user-base (i.e., MCR is therapy target for children with autism [11]), *Go Go Games* has the potential to expand the reach of how many children receive treatment for overselectivity, despite its limitations.

Additionally, we demonstrated effectiveness of *Go Go Games* in teaching MCR to a subset of the population that needs support. The participants who struggled with MCR in their baseline assessment and received the full treatment dose showed improvements in their final assessment. This study has important implications for training this specific skill using this specific game and for training language skills using digital therapies more generally. Given that teaching MCR to reduce overselectivity could potential support the high proportion of preliterate AAC users in more effectively accessing their AAC systems, ongoing research is warranted to understand the role of overselectivity and how to build smart systems to reduce its impact.

People with severe disabilities often exhibit secondary disabilities [21], thus, supporting these circumstances may be essential to the effectiveness of the tool. One way to merge the specialized supports required of these complex conditions would be to train designers of assistive technology for preliterate users in methods to reduce overselectivity, as overselectivity may be a potential co-occurring condition. Alternatively, we could address these complex needs through collaboration across several domains of disability to develop a shared vision of the multitude of needs to address in an assistive technology aimed at a pivotal cognitive skill. This collaborative effort could result in a decision tree for assistive technology designers to concurrently support the unique cognitive, physical, and behavioral as well as social needs of AAC users.

For those relying on AAC devices but struggling with MCR, game mechanisms to reduce overselectivity is paramount. For example, designers of therapeutic games could support a more diverse user base through scaffolding and additional supports and/or build a screener via an experiential tutorial to confirm the user has the prerequisite skills to benefit from the system. A screener could address the wide range of abilities and diverse sets of disabilities (e.g., test various configuration of audio, visual, and motor interactions or even configure itself intelligently based caregiver feedback and user interaction). In this way, input could be varied depending on a user's motor skills, interactions could be varied depending on cognitive skills, and so on as is suggested in ability-based design [32]. For example, P1 who struggles to swipe an image on the iPad may be screened for overselectivity using a variety of interactions (e.g., variations ranging from soft quick tap to sustained drag).

Likewise, system interfaces should be adjustable. For example, the background of a computer game, which can be aesthetically pleasing or a key element to provide enjoyment for some children, can be distracting for children who struggle with overselectivity. A variety of options should be made available during both implementation and game play for customizing these types of non-essential game elements. For example, simply making the game play window smaller limits the visual field the child must interpret. Similarly, being able to toggle on and off various background elements or object size could provide some customizability for children with special needs. Lastly, the mechanics of the reward schedule could be customizable to increase tolerance to in-game feedback. These kinds of solutions are essential for therapeutic games to meet the specific needs of the children who are engaged in them and could be configured by a trained specialist or a parent. However, this kind of customization may also be necessary and useful for a wider variety of games to make them accessible to more children for entertainment or educational purposes.

7. CONCLUSION

In this paper, we presented initial findings toward a body of evidence of the effectiveness of a mobile behavioral intervention in the form of an iPad game for minimally verbal children who struggle with overselectivity. Specifically, single-case design results suggest the mean scores shifted positively for three of six participants after nine days of intervention. Additionally, we describe the conditions under which the system worked well and those areas that require additional design work. We presented generalizable technological design features drawn from evidenced-based therapies to consider in the design of future assistive technologies for people who are preliterate and minimally verbal.

These results leave open research questions regarding the potential for customizing the settings related to vision, motor movements, attention spans, and feedback tolerance for a diversity of issues that challenged users with overselectivity. These design challenges are not trivial, and they represent an important scope of challenge for making technology and games accessible. These systems have the potential to provide both therapeutic benefit and entertainment value. However, this value increasingly relies on the ability to adapt to a variety of users and contexts.

Finally, to truly understand the potential power of these technologies and the barriers to their use, a large evidence-base must be assembled. This work should be followed up with collaborative efforts across a variety of researchers who support preliterate AAC users, as well as evaluation studies that are longer, and therefore provide higher dosages of the game to the children, and include additional participants with other co-occurring challenges to understand the full scope of potential impact and engagement.

ACKNO WLEDGMENTS

We thank the school, staff, children and parents who consented to have their children participated in this work. We thank the behavior therapist, Melissa Kramer, for the translational work that provided a smooth study. We also thank the research assistants, Andrea Conejo-Toledo, and Koramya Arriaga who assisted in the training materials for data collection. We thank STAR lab for early input on this manuscript. We thank Robert and Barbara Kleist for funds to support this work. This research was approved by human subjects protocol #2015-1871.

REFERENCES

[1] Meghan Allen, Joanna McGrenere, and Barbara Purves. 2007. The design and field evaluation of PhotoTalk: a digital image communication application for people. *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*, ACM, 187–194.

[2] Rolf Black. 2011. The Phonicstick: a joystick to generate novel words using phonics. *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, ACM, 325–326.

[3] John O. Cooper, Timothy E. Heron, William L. Heward, and others. 2007. Applied behavior analysis.

[4] Chata A. Dickson, Sharon S. Wang, Kristin M. Lombard, and William V. Dube. 2006. Overselective stimulus control in residential school students with intellectual disabilities. *Research in Developmental Disabilities* 27, 6, 618–631. http://doi.org/10.1016/j.ridd.2005.07.004

[5] William V. Dube and Krista M. Wilkinson. 2014. The Potential Influence of "Stimulus Overselectivity" in AAC: Information from Eye-tracking and Behavioral Studies of Attention. *Augmentative and alternative communication (Baltimore, Md. : 1985)* 30, 2, 172–185. http://doi.org/10.3109/07434618.2014.904924

[6] Gene S Fisch. 2001. Evaluating data from behavioral analysis: visual inspection or statistical models? *Behavioural Processes* 54, 1–3, 137–154. http://doi.org/10.1016/S0376-6357(01)00155-3

[7] Alexis Hiniker, Joy Wong Daniels, and Heidi Williamson. 2013. Go go games: therapeutic video games for children with autism spectrum disorders. *Proceedings of the 12th International Conference on Interaction Design and Children*, ACM, 463–466. Retrieved September 24, 2014 from http://dl.acm.org/citation.cfm?id=2485808

[8] Alexis Hiniker, Hyewon Suh, Sabina Cao, and Julie A. Kientz. 2016. Screen Time Tantrums: How Families Manage Screen Media Experiences for Toddlers and Preschoolers. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 648–660. http://doi.org/10.1145/2858036.2858278

[9] Robert H. Horner, Edward G. Carr, James Halle, Gail McGee, Samuel Odom, and Mark Wolery. 2005. The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education. *Exceptional Children* 71, 2, 165–179. http://doi.org/10.1177/001440290507100203

[10] Amie M King, Kathryn W Brady, and Grayce Voreis. 2017. "It's a blessing and a curse": Perspectives on tablet use in children with autism spectrum disorder. *Autism & Developmental Language Impairments* 2, 2396941516683183. http://doi.org/10.1177/2396941516683183

[11] Lynn Kern Koegel, Robert L. Koegel, Joshua K. Harrower, and Cynthia Marie Carter. 1999. Pivotal Response Intervention I: Overview of Approach. *Journal of the Association for Persons with Severe Handicaps* 24, 3, 174–185. http://doi.org/10.2511/rpsd.24.3.174

[12] Robert L. |Kern Koegel Koegel. 2006. *Pivotal Response Treatments for Autism: Communication, Social, and Academic Development*. Brookes Publishing Company.

[13] David McNaughton and Janice Light. 2013. The iPad and Mobile Technology Revolution: Benefits and Challenges for Individuals who require Augmentative and Alternative Communication. *Augmentative and Alternative Communication* 29, 2, 107–116. http://doi.org/10.3109/07434618.2013.784930

[14] Bart Michiels, Mieke Heyvaert, Ann Meulders, and Patrick Onghena. 2016. Confidence intervals for single-case effect size measures based on randomization test inversion. *Behavior Research Methods*. http://doi.org/10.3758/s13428-016-0714-4

[15] Pat Mirenda. 2009. Promising Innovations in AAC for Individuals With Autism Spectrum Disorders. *SIG 12 Perspectives on Augmentative and Alternative Communication* 18, 4, 112–113. http://doi.org/10.1044/aac18.4.112

[16] Fereshteh Mohammadzaheri, Lynn Kern Koegel, Mohammad Rezaei, and Enayatolah Bakhshi. 2015. A randomized clinical trial comparison between pivotal response treatment (PRT) and adult-driven applied behavior analysis (ABA) intervention on disruptive behaviors in public school children with autism. *Journal of autism and developmental disorders* 45, 9, 2899–2907.

[17] Samuel L. Odom, Ellen Brantlinger, Russell Gersten, Robert H. Horner, Bruce Thompson, and Karen R. Harris. 2005. Research in Special Education: Scientific Methods and Evidence-Based Practices. *Exceptional Children* 71, 2, 137–148. http://doi.org/10.1177/001440290507100201

[18] Katie O'Leary, Charles Delahunt, Patricia Dowden, et al. 2012. Design goals for a system for enhancing AAC with personalized video. *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*, ACM, 223–224.

[19] K. Phelan and H. E. McDermid. 2011. The 22q13.3 Deletion Syndrome (Phelan-McDermid Syndrome). *Molecular Syndromology* 2, 3–5, 186–201. http://doi.org/10.1159/000334260

[20] Bertram O. Ploog. 2010. Stimulus Overselectivity Four Decades Later: A Review of the Literature and Its Implications for Current Research in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* 40, 11, 1332–1349. http://doi.org/10.1007/s10803-010-0990-2

[21] Jenny Preece, Helen Sharp, and Yvonne Rogers. 2015. Interaction Design: Beyond Human-Computer Interaction. Second Edition. *MyScienceWork*.

[22] Deborah Preston and Mark Carter. 2009. A Review of the Efficacy of the Picture Exchange Communication System Intervention. *Journal of Autism and Developmental Disorders* 39, 10, 1471–1486. http://doi.org/10.1007/s10803-009-0763-y

[23] Suzanne Prior, Annalu Waller, and Thilo Kroll. The CHAMPION software project. *Proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, ACM, 287–288.

[24] Sarah R. Rieth, Aubyn C. Stahmer, Jessica Suhrheinrich, and Laura Schreibman. 2015. Examination of the prevalence of stimulus overselectivity in children with ASD. *Journal of Applied Behavior Analysis* 48, 1, 71–84. http://doi.org/10.1002/jaba.165

[25] Henry S. Roane, Wayne W. Fisher, Michael E. Kelley, Joanna L. Mevers, and Kelly J. Bouxsein. 2013. USING MODIFIED VISUAL-INSPECTION CRITERIA TO INTERPRET FUNCTIONAL ANALYSIS OUTCOMES: FUNCTIONAL ANALYSIS INTERPRETATION. *Journal of Applied Behavior Analysis* 46, 1, 130–146. http://doi.org/10.1002/jaba.13

[26] Laura Schreibman. 1997. The study of stimulus control in autism. *Environment and Behavior*, 203–209.

[27] Laura Schreibman and Robert L. Koegel. 1982. Multiple-cue responding in autistic children. *Advances in Child Behavioral Analysis & Therapy* 2, 81–99.

[28] Shari Trewin, Cal Swart, and Donna Pettick. 2013. Physical Accessibility of Touchscreen Smartphones. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 19:1–19:8. http://doi.org/10.1145/2513383.2513446

[29] Keith Vertanen. 2013. A collection of conversational AAC-like communications. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 31.

[30] Tonio Wandmacher, Jean-Yves Antoine, and Franck Poirier. 2007. SIBYLLE: a system for alternative communication adapting to the context and its user. *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*, ACM, 203–210.

[31] Karl Wiegand and Rupal Patel. 2012. SymbolPath: a continuous motion overlay module for icon-based assistive communication. *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*, ACM, 209–210.

[32] Jacob O. Wobbrock, Shaun K. Kane, Krzysztof Z. Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-Based Design: Concept, Principles and Examples. *ACM Trans. Access. Comput.* 3, 3, 9:1–9:27. http://doi.org/10.1145/1952383.1952384