Photocopy and Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature ____________________________________

Date _______________________________________

The Effects of Environmental Sound and Animation on the Naming Accuracy of Graphic Symbols in Children

by

Sarah Aldrich

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in the Department of Communication Sciences and Disorders

Idaho State University

Spring 2021

Committee Approval

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Sarah Aldrich find it satisfactory and recommend that it be accepted.

Dr. Kris Brock,
Major Advisor

Dr. Diane Ogiela,
Committee Member

Ms. Mary Van Donsel,
Committee Member

Dr. Jenn Gallup,
Committee Member

Dr. Amanda Eller,
Graduate Faculty Representative

Human Subjects Committee Approval

April 17, 2020
Kristofer Brock
College of Rehabilitation Comm. Sciences

RE: IRB-FY2019-224 The Effects of Environmental Sound and Animation on the Naming Accuracy of Graphic Symbols for Children with Developmental Disabilities

Dear Dr. Brock:

Thank you for your responses to a previous full-board review of the study listed above. These responses are eligible for expedited review under OHRP (DHHS) and FDA guidelines.

You are granted permission to continue your study as described effective immediately. The study is next subject to continuing review on or before April 17, 2021, unless closed before that date

Please note that any changes to the study as approved must be promptly reported and approved. Some changes may be approved by expedited review; others require full board review. Contact Tom Bailey (208-282-2179; fax 208-282-4723; email: humsubj@isu.edu) if you have any questions or require further information.

Sincerely,

Ralph Baergen, PhD, MPH, CIP Human Subjects Chair

Table of Contents

List of Tables

The Effects of Environmental Sound and Animation on the Naming Accuracy of Graphic Symbols for Children with Developmental Disabilities

Thesis Abstract--Idaho State University (2021)

Considerable enthusiasm is being generated for the inclusion of animation and sound as a teaching tool for children with complex communication needs. However, the effects of environmental sound within animated verb symbols on naming accuracy is confined to a single study. The purpose of this study was to examine the impact of animated verb graphic symbols with embedded sound on naming accuracy and response time. Participants included 21 children with typical development between the ages of 4;0 to 4;11 (years; months) and 6 children with developmental disabilities between the ages of 8;1 to 15;0. Results indicated that environmental sounds did not significantly enhance the naming accuracy of animated verb symbols or response time. Implications for clinical practice and future research will be discussed.

Keywords: Augmentative and alternative communication; Environmental sounds; Graphic symbols; Naming; Response time

Chapter 1: The Effects of Environmental Sound and Animation on the Naming Accuracy of Graphic Symbols with Developmental Disabilities

Approximately 7.6% of children with disabilities, such as autism spectrum disorder (ASD), have a communication impairment so severe that they require an augmentative and alternative communication (AAC) system to replace or supplement natural speech and language (Keeny & Koogan, 2011). AAC has been defined as an area of practice that "... compensate(s) for temporary or permanent impairments, activity limitations, and participation restrictions ..." (ASHA, 2005, para. 1) by incorporating "... tools and strategies [such as symbols, pictures, and speech-generating devices] that an individual uses to solve every day communicative challenges" (ISAAC, 2017, para. 1). To meet these challenges, the foundation of AAC includes the symbol, and specific to this study is the graphic symbol. Graphic symbols fall on a spectrum between transparent (i.e., readily understood symbol-referent relationship) and opaque (i.e., no relationship between the symbol and its referent). Additionally, there are two forms of graphic symbol, either static or animated and each can include sound to enhance the symbol's transparency.

While research in animation has increased in the last decade, the utility of animation for communication purposes has focused on increasing the iconicity of verbs and prepositions and reducing the cognitive demand placed on users to find and name symbols (Mineo et al., 2008; Fujisawa et al., 2011; Lee & Hong, 2013; Schlosser et al., 2011; 2012; 2014; 2019). Additionally, a single study investigated the effects of animated symbols with and without sound using 3-year-old children without disability (Harmon et al., 2014). Overall, these studies indicated that children named and identified verbs and prepositions with higher accuracy in the animated *and* animated sound conditions compared to traditional static symbol and no sound conditions.

Research into the effects of animation and sound are slowly progressing, but generalization of previous outcomes to the broader population of individuals with complex communication needs is lacking. Specific to the current study, Harmon et al. (2014) only included 3-year-old children without disability in their study. They cannot rule out any age effects because children of a single age were included. It is possible that 4-year-olds without disability no longer require the additional audio support embedded in animated graphic symbols to name verbs. Therefore, it is important to study the effects of animation and sound on communication outcomes in various populations.

Evidence Base for Animation and Sound

Animation

Animation is defined as a series of sequential frames that change slightly over time to produce a specific movement that explains a dynamic construct (Bétrancourt & Tversky, 2000). Generally, animation serves four purposes: (a) animation attracts a learner's attention; (b) conveys information related to a process; (c) portrays the completion of a procedure; and (d) demonstrates changes over time (Berney & Bétrancourt, 2016). However, several moderating factors were found to influence learning from animations (Berney & Bétrancourt, 2016; Höffler & Leutner, 2007). The most noteworthy moderating factors for AAC purposes include the abstraction quality of animation, the pace of animation, and the inclusion of sound.

As Ploetzner and Lowe (2012) discussed, abstraction refers to how closely a graphic representation resembles the object or construct being represented, and this principle is akin to the iconicity principle used within the AAC literature (Bloomberg et al., 1990). Specifically, there is an inverse relationship between iconicity of a symbol and the difficulty associated with learning the symbol (Angermeir et al., 2008). In addition to abstraction, Berney and Bétrancourt (2016) noted that the pace of the animation must be conducive to learning, that is, slow enough to engage and learn from the content. Finally, sound...Overall, research across multiple disciplines has found that animation is facilitative when implemented correctly (see Berney & Bétrancourt, 2016). People interact within a multisensory environment that provides constant input, and their cognitive systems are capable of appropriately processing each sensory input to interpret the experience (Driver & Spence, 2000).

Animation Evidence in AAC

In a landmark study, Schlosser et al. (2014) randomly assigned 220 typically developing preschoolers across three age groups (3-, 4-, and 5-year-old) to combinations of symbol set (ALP Animated Graphics Set or Picture Communication Symbols [PCS]), symbol format (animated or static), and word class (verbs or prepositions). Results indicated that animated verbs were more readily named than static verbs. Additionally, the ALP symbols were named more readily than the PCS counterparts. Schlosser et al. (2019) replicated parts of their previous results in children with autism by investigating the effects of symbol format on identification accuracy of ALP verb graphics. Results indicated that the children identified a significantly greater number of animated graphic symbols when compared to static graphic symbols. Overall, animation has a promising role in AAC intervention, but further research is needed.

Sound and AAC

Auditory research as well as AAC specific guidance suggests that sounds may be beneficial to individuals with complex communication needs (Harmon, et al., 2014; Ogletree et al., 2018; Sussman, 2017). Each environment has a myriad of sounds from background conversations to traffic, and the auditory system is responsible for discerning which sounds require our attention and which can go unattended. This method, called auditory scene analysis, is a multilevel, automatic stimulus-driven procedure where we process and have access to all of the sounds in our environment (Sussman, 2017). The important sounds are attended to (e.g., someone calling your name or a warning siren), but the unattended sounds are still processed and maintained in memory for later use should they become important (Sussman, 2017).

The most important variables for sound in AAC are the concepts of congruence, redundancy, and nomic mapping. First, semantic congruence refers to coupling a visual input with an auditory input (Chen & Spence, 2010; Nava et al., 2016). When visual and sound stimuli are congruent (e.g., bacon sizzling in a pan), the brain better integrates that information for use (Koppen, Alsius, & Spence, 2008). Harmon et al., (2014) hypothesized that semantic congruence could be beneficial for individuals with complex communication needs because it provides an additional avenue for learning. Second, the connection between a specific sound and the visual stimuli can be either redundant or non-redundant (Nava et al., 2016). Redundancy refers to crossmodal sensory input or when an aspect of a particular stimulus produces sensory input for more than one modality. An example includes clapping hands, where the movement is seen and the sound processed. However, as Harmon et al. (2014) pointed out, the strength of the auditory-symbol association can be influenced by how closely the sound relates to the symbol (i.e., sound iconicity). The third and final variable, nomic mapping (Gaver, 1986), occurs when the sound source and the sound are learned associations (e.g., cows mooing). Nomic mappings lend themselves well to verbs secondary to the obvious links between the sound and the action (e.g., Gaver, 1986). Overall, pairing environmental sounds with animated graphic symbols may enhance the naming of verb symbols in children with and without complex communication needs.

Finally, Harmon et al. (2014) suggested that Paivio's (1986) dual coding theory, where information is processed by a linguistic system and a nonverbal system, was partially responsible for their findings. That is, graphic symbols accompanied by auditory input are coded in the nonverbal system twice, once for the sound and once for the graphic symbol (Harmon et al., 2014). While seminal, the research only included 3-year-olds with typical development. Moreover, younger children may benefit from the additional auditory support whereas older children may not see the same facilitative effects on naming.

In sum, previous research indicates that a majority of children with typical development and children with various developmental disabilities are better able to name or identify verbs represented as animated graphic symbols (Fujisawa et al, 2011; Mineo et al., 2008; Schlosser et al., 2012; 2014; 2019). However, very little is known about the effects of animation and sound on graphic symbol naming in older children with typical development and children with complex communication needs. Therefore, the purpose of this study was two-fold. First, it is important to replicate the research of Harmon et al. (2014) and determine if environmental sounds paired with animated graphic symbols enhance verb naming accuracy in 4-year-old children with typical development and children with developmental disabilities (ASD, Down Syndrome, or intellectual disability) between the ages of 5 and 15 years of age. Second, the verb naming errors will be analyzed to identify any patterns in an attempt to improve the iconicity of the symbols used in this study.

Chapter 2: Method

Participants

Twenty 4-year-old children with typical development and five elementary school-age children between the ages of 8;0 and 11;11 with developmental disabilities participated in this study. Participants were recruited from local school districts, the university clinic, and private clinics in the Boise metropolitan area and Pocatello semi-metropolitan area. The 4-year-old participants (see Table 1) met the following criteria: (a) English as the primary language spoken in the home; (b) no uncorrected vision or hearing impairments; and c) expressive or receptive knowledge of the target verbs used in the study per a pre-assessment screening or, if noncompliance was observed, based on parent confirmation.

The participants with developmental disabilities (see Table 1) met the following inclusion criteria: (a) English as the primary language spoken in the home; (b) no uncorrected vision or hearing impairments; (c) an unequivocal diagnosis of autism spectrum disorder, Down syndrome, or an intellectual disability as indicated by parent records, school records, or speech-language pathologist records; (d) scores resulting in > 1.5 standard deviations below the mean within two or more domains (i.e., Communication, Daily Living Skills, Socialization, and Motor Skills) on the *Vineland Adaptive Behavior Scales-Third Edition* (Sparrow et al., 2016); (e) mild to severe autism severity as indicated by the *Childhood Autism Rating Scale-2* (*CARS-2*; Schopler et al., 2010); (f) expressive or receptive knowledge of the target verbs used in the study per a pre-assessment screening or, if noncompliance was observed, parent report of understanding; and (g) free from sensory and motoric deficits that could impact the speech mechanism as indicated by medical, school, or SLP records.

Settings

Because this project was interrupted by COVID-19, some of the participants completed the study under face-to-face conditions while others completed the study through video conferencing software (see Table 1). For those that completed the study in-person, the tasks were completed in quiet rooms within the participants' home, the site for an AAC camp, daycare premises, or university clinic. The rooms were accessible to parents or counsellors at any time. The children who participated through video conferencing were in their own homes with a caregiver seated next to them.

Experimenters

The experimenters included a certified speech-language pathologist researcher and graduate student research assistant.

Table 1

Participant Demographic and Assessment Information

Participant Type*	Gender	Age in Months	Expressive Verb Screening	Receptive Verb Screening	VABS-3 Score for Communication	VABS-3 Score for Daily Living Skills	VABS-3 Score for Socialization	CARS-2 Score
Developmental Disability	2 F 4 M	138	58.3 (29.3)	96.3 (5.7)	45.4 (36.8)	54.4 (31.9)	51.0 (20.4)	43.3(11.6)**
Typical Development	13 F 8 M	52.4	62.5 (12.3)	100 (0)	93.3 (13.4)	96.2 (8.6)	97.3 (9.3)	NA

Note. * = Two children with disabilities and six children with typical development took part in this study through Zoom, while the remainder were face-to-face. ** = Four children had a diagnosis of autism, one child had an intellectual disability, and one child had cerebral palsy.

Materials

Verbs

The 18 verbs used in this study were the same verbs used by Harmon et al. (2014): BREAK, CATCH, CLAP, CLOSE, CRY, DRAW, DRY, HIT, KISS, LICK, OPEN, POUR, SHAKE, SPILL, SPLASH, SWIM, SWEEP, and WIPE. These verbs were selected from the MacArthur-Bates Communication Development Inventory (CDI; Fenson et al., 1994), in which half of the norming population acquired by 30 months of age. Sixteen (16) of the verbs are transitive, meaning that they take a direct object. This leaves two intransitive verbs (i.e., cry and draw) where a direct object was not required. The greater number of transitive verbs used in this study was based on previous research. Schlosser et al. (2011) found that these 18 verb symbols were not readily nameable (i.e., < 70%), thereby justifying the addition of sound to potentially enhance naming.

Graphic Symbols

The ALP animated graphics set were chosen for this study because the purpose of this paper is to replicate the findings from Harmon et al (2014). Additionally, this study exclusively included symbols that had action central sounds (i.e., produced directly from the action [*crying])* rather than action peripheral sounds (i.e., related sounds that do not occur from the action [*drive*]). This is important because action peripherals may be subjective to the receiver, potentially influencing the results. For example, individuals may associate a motor with driving while others envision angry motorists in traffic.

Sounds, Recording, and Speakers

The first author acquired the two sets of environmental sound stimuli from Harmon et al. (2014) for this replication study. For the experimental task, 18 sounds were previously recorded using an Olympus Digital Voice Recorder VN-8100PC ™ with an Olympus Electret Condenser Microphone ME51S ™ attachment in certain cases within a soundproof room. Additionally, two sounds (*CLOSE* and *CRY*) included a high-quality recording from the Sound Ideas Series 6000 General Sound Effects Library (http://www.sound-ideas.com/soundeffects/eries-6000-sound-effects-library.html). All sounds were determined to have an Fs of 48000 and 16 nbits and presented through the computer's internal speakers. Finally, the experimenter recorded each session using a Sony CX440 HandyCam during face-to-face interactions for procedural integrity and reliability purposes. Once COVID-19 started, the experimenter screen recorded each interaction while the caregiver used their smartphone to record each interaction from their home. This caregiver recording allowed for procedural integrity and intra-rater reliability analysis as well as response latency analysis.

Hardware and Software

The experiment required a Dell Inspiron laptop computer with a 15-inch display, i7-8550U Processor with 16GB of memory, and a 512GB solid-state drive. The Dell laptop was selected because the experimental task files were large secondary to the video and audio stimuli employed. These computer features prevented lag and freezing, which could affect naming and response latency outcomes.

Prior to COVID-19, participants engaged in the experiment through E-Prime-3.0® (Psychology Software Tools, 2020). E-Prime software was selected for this study because it automatically (a) presented the experimental stimuli and (b) measured participant response accuracy and latencies. This software provided automated, objective data that eliminated inconsistencies related to using multiple stopwatches while mitigating experimenter bias and error.

During COVID-19, participants engaged in the experiment through video conferencing software and were shown a PowerPoint slide show presentation equivalent to the E-Prime presentation. That is, every experimental setting and associated time (see Procedures section) were the same in PowerPoint and E-Prime. A digital text file allowed the experimenter to track participant responses. Additionally, two digital recordings captured the entire experiment. The first recording was the experimenter's screen recording of the video conference. This screen recording was simply a back-up recording and never had to be used in any analysis. The second recording included the caregivers' smartphone to record the child engaged with the computer/experiment and their verbal responses. The caregivers' recording was uploaded to the video editing software, Camtasia®, to calculate response latency, or the time between stimuli presentation and verb naming (see Procedures for details). The caregivers' recording was selected for response latency analysis because the authors could not rule out lag time as the experiment "travelled" to the participants' computer screen. Additionally, using the caregivers' video ruled out lag associated with slower Wi-Fi connections.

Finally, four of children with a developmental disability could not use natural speech production to engage in the task. Therefore, these individuals used their personal speech-generating device (SGD) or the Proloquo2Go app on a tablet. Regardless of the system, the experimenter and caregiver randomly inserted 18 graphic symbols into a separate communication folder. These 18 symbols corresponded to the 18 verbs used in the study. These participants followed the same procedures except that their naming was the selection of one of the 18 symbols programmed into their device.

Research and Design

This study used a within-subjects design. Presentation of the experimental conditions was counterbalanced across participants. After 1 to 3 weeks, participants engaged in the second half of the experiment. This controlled for practice effects while also mitigating developmental effects. Finally, the institutional review board approved this research

Procedures

Pre-assessment

In one 90-minute session, three pre-assessments were conducted: (a) an expressive knowledge of verbs screening task of the target verbs used in the study, (b) a receptive knowledge of verbs screening task, (c) the *VABS-3* and (d) the Child Autism Severity Rating Scale -2 (Schopler, et al., 2010). These tasks were completed face-to-face or through video conferencing software.

During session 1, a randomized expressive screening of verbs was conducted to ensure word knowledge of the verbs used in the study. This was necessary because the experiment is investigating the ability of children to process and name symbols representing verbs already in their vocabulary. First, the experimenter performed an action with a standard set of props. Then the child, when asked, named the action within 10 seconds. If there was no response after 10 seconds, the experimenter repeated the action. If no response was recorded again, the trial was marked incorrect. For those incorrect trials, a randomized receptive task was administered. The experimenter named an action and asked the child to perform that action using a standard set of props. For example, "*Can you BOUNCE the ball?*" If the child did not bounce the ball in 10 seconds, the question was repeated. If no response, the child was given an indirect verbal prompt, "*I bet the ball can go high. Can you bounce the ball?*" and another 10 seconds to respond. In the event that a child did not demonstrate understanding of a verb either the expressive or receptive task, the experimenter asked the parent if the child has demonstrated understanding of the verb in natural settings. If the parent responded in the affirmative the child was allowed to continue with the study.

Familiarization Task

First, the face-to-face familiarization task included the use of E-Prime-3.0. Second, participants using video conference were administered the same familiarization task using Microsoft PowerPoint©. PowerPoint was used because E-Prime locked the experimenter's ability to engage through video conferencing software reliably. It is important to note that each software program had the same timing and stimuli of the same size and image quality.

To start the familiarization trials, a digitized female voice in both software programs said, "*Hi kiddo! Let's play a guessing game on the computer. Are you ready?* If the child responded affirmatively, the experimenter clicked on the spacebar to indicate child assent If the child responded in the negative they were given an opportunity to ask questions. The next slide appeared and the digitized voice said, "*You'll see lots of short movies on the computer. Then the computer will tell you to name the movie. Make sure you name the movie before the red screen pops up."* Then, a third screen appeared, and the recording said, *First, I will show you how to play the game. Here we go!"* The child was presented with a green screen lasting for 5.5 s, and the digitized voice said, *"Look at the computer…What's this?"* Next, a single animated graphic symbol with or without sound appeared. The participant was expected to name the target symbol within 30 s. After 30 s, a red screen automatically appeared to signal the end of the trial. If the participant was correct, the experimenter read the script, *Yes, that is [verb]*. If incorrect (2), the experimenter said, *No, that is not [incorrect word]. That is [correct verb]*." After this feedback (approximately 5 seconds), the experimenter asked, "*Is the sound loud enough?*" The volume was adjusted as necessary. The red screen was engaged as long as necessary; however, by clicking the spacebar, the next trial was administered. This was done to insure the participant understood the feedback and the task. The familiarization trials were repeated until the task was completed with 100% naming accuracy.

Experimental Task

The experimental task followed the same structure as the familiarization task. However, in this task, all 18 verbs were presented, and a 5 s inter-trial interval was provided between the red screen and the subsequent green screen. The experimenter provided neutral feedback (e.g., *"Keep up the god work!"*) during the inter-trial interval. During E-Prime administration, the experimenter recorded the participants' naming responses and response times manually using a Bluetooth keyboard connected to the experimental computer. The keyboard was hidden behind the experiment computer's display. The keys "1" and "2" served two purposes: (a) to record naming accuracy (1 correct; 2 incorrect) and (b) record response latency. To ensure reliable response latency recording, the experimenter placed an index finger on the letter "Q" just below the "1" and "2" keys. Thus, the distance to those recording keys remained the same throughout the experiment and did not affect response times.

The automatic naming accuracy and response time data tracking procedure afforded by E-Prime was changed to accommodate the video conferencing software and PowerPoint experimental task administration (see Appendix A for video conference procedures). First, the experimenter screen recorded the video conference call while the caregiver used their smartphone to record their child and their computer screen. Second, the experimenter manually recorded data and marked "+" for a correct naming response or an "X" for an incorrect naming response on an electronic, text-based data sheet. Second, participants saw a green screen for 5.5 *s* and then heard an audible beep along with the presentation of the animated verb, signaling the start of the experimental trial that lasted 30 *s*. The audible beep allowed the experimenter to measure response latency (i.e., time between the audible beep and verb naming) through the caregivers' digital recording.

Regardless of experimental software, only the first intelligibly spoken word, as judged by the researchers, was used for data analysis purposes to ensure consistency across participants. Spoken words deemed to be approximations were flagged and reviewed using the recordings. Video time stamps were then used to re-evaluate response times and subsequently confirmed by two researchers. A "no response" was automatically recorded as a 30 s response time and thrown out of subsequent *t*-test analysis.

Data Analysis

First, the descriptive statistics for the experimental tasks were calculated. Additionally, two dependent t-tests were conducted: (1) verb naming accuracy between the sound and no sound conditions and (2) response time (ms) differences between the two conditions. Response naming accuracy was defined as either the production of the intended verb name or an approved synonym (e.g., hop for jump). In the E-Prime software experiment, response latency was defined as the time between the initial presentation of the stimuli and the end of the verb name spoken aloud (i.e., when the experimenter selected the "2" key). For those who engaged with the PowerPoint software, response time was defined as the elapsed time between the audible beep signaling the start of the trial and the participant speaking the name of the verb. To calculate this difference in time, the caregivers' digital recordings were inserted into the video editing software. The experimenter found the audible beep and the participants' spoken word in the audio component of the file. Only correct responses were used in response latency analyses.

Second, we applied the heuristic criteria from Schlosser et al. (2011) to assess how well our symbols performed in this study. Specifically, symbols in this study were considered satisfactory if participants named the symbol with > 70% accuracy. This measure deemed within word class synonyms (e.g., toss for throw) as correct. For name agreement, participants must use the exact verb as indicated by the research team with > 60% accuracy. Anything below that criterion was deemed to be underperforming. To extend Schlosser et al.'s (2011) work, a response time heuristic was created post-hoc to determine how quickly a symbol was named. This heuristic provides practicing professionals with an objective metric to determine if a symbol is transparent enough for efficient naming purposes. In addition to these variable-specific criteria, Schlosser et al.'s (2011) across-variable heuristic was used to classify symbols as performing exceptionally, effectively, adequately, or inadequately. Symbols performing exceptionally met the three symbol performance criteria as discussed above (i.e., naming, name agreement, and response time). Symbols performing effectively met two of the three criteria, adequately performing symbols met one of the three criteria, and inadequately performing symbols did not meet any of the criteria. This heuristic provides practicing professionals with an objective metric to determine if a symbol is being processed efficiently for naming purposes.

Finally, descriptive error analysis was completed using the categories and definitions from Harmon et al. (2014) which were adapted from Masterson, Druks, and Gallienne (2008). The categories included (a) semantic errors, (b) visual errors, (c) other errors (e.g., word class errors, mixed errors), and (d) auditory errors. The list of categories and their definitions can be found in Appendix B.

Inter-Observer Agreement

Using the video recordings, inter-observer agreement data were collected for 30% of the participants. A research assistant, blinded to the purpose of the study, served as the second rater. Data for both the pre-assessment task and the experimental task were compared. The percent agreement was calculated by dividing the number of items that were recorded the same way by the total number of data points collected on a data sheet and multiplied by 100. Disagreements resulted in a meeting of two of the researchers to reach consensus. If consensus was not reached, a third researcher broke the tie. Inter-observer agreement was 98.4% for naming accuracy, 100% for response latency, and 88.89% for error pattern analysis. The researchers met multiple times to review error patterns analysis and came to consensus reaching 100% agreement.

Procedural Integrity

Procedural integrity calculations were not necessary for several of the participants prior to COVID-19. This is because E-Prime software automated all procedures including inter-trial intervals, neutral verbal reinforcement, and auditory directions. Therefore, procedural integrity was 100%. However, E-Prime was not compatible with Zoom during COVID-19. Therefore, as indicated above, four participants were administered the experiment using PowerPoint®. Procedural integrity data were calculated for 20% of the participants who engaged with the PowerPoint and resulted in 100% accuracy.

Chapter 3: Results

Naming Accuracy and Response Latency

For 4-year-old children with typical development, the *t*-test revealed no significant naming accuracy difference ($t(20) = .18$, $p = 0.86$, Cohen's $d = .04$) between the animated symbols with environmental sounds ($M = 74.60$, $SD = 18.72$) and the animated graphic symbols without sound ($M = 74.07$, $SD = 14.09$). For children with developmental disabilities, the *t*-test revealed no significant naming accuracy difference ($t(5) = -1.37$, p = 0..23, Cohen's $d = .56$) between the animated symbols with environmental sounds ($M = 75.93$, $SD = 21.3$) and the animated graphic symbols without sound ($M = 70.37$, $SD = 18.1$).

Table 2 summarizes the percentage of correct responses by group for each symbol in each condition. The children with typical development performed better on eight of the symbols in the sound condition (*KISS*, *SPLASH*, *DRAW*, *SWIM*, *POUR*, *HIT*, *LICK*, and *WIPE*) and on seven of the symbols in the no sound condition (*CLOSE*, *SHAKE*, *BREAK*, *CATCH*, *CRY*, *CLAP* and *SPILL*). Three symbols (*DRY*, *OPEN*, and *SWEEP*) performed similarly regardless of condition. For 5 of the 18 verbs, children with disabilities performed better with sound than without sound (*KISS, SHAKE, LICK, HIT and CRY*); however, on 1 of the symbols (*SPILL*) they performed better without sound. The differences in performance on the remaining 12 symbols were negligible regardless of the condition.

Table 2

Experimental Verb Naming Accuracy for all Participants

Symbol	Participant	Sound % (SD)	No Sound % (SD)
Kiss	Disability	100.0 (0.0)	83.33 (40.8)
	No Disability	95.2 (21.8)	76.2 (43.6)

Splash	Disability	66.7 (51.6)	66.7 (51.6)
	No Disability	57.1 (50.7)	52.4 (51.2)
Draw	Disability	83.3 (40.8)	83.3 (40.8)
	No Disability	95.2 (21.8)	90.5 (30.1)
Open	Disability	83.3 (40.8)	83.3 (40.8)
	No Disability	71.4 (46.3)	71.4 (46.3)
Close	Disability	83.3 (40.8)	83.3 (40.8)
	No Disability	71.4 (46.3)	85.7 (35.9)
Swim	Disability	100.0 (0.0)	100.0 (0.0)
	No Disability	100.0 (0.0)	85.7 (35.9)
Shake	Disability	66.7 (51.6)	50.0 (54.8)
	No Disability	71.4 (46.3)	81.0 (40.2)
Sweep	Disability	100.0 (0.0)	100.0 (0.0)
	No Disability	90.5 (30.1)	90.5 (30.1)
Hit	Disability	50.0 (54.8)	16.7 (40.8)
	No Disability	71.4 (46.3)	57.1 (50.7)
Break	Disability	50.0 (54.8)	50.0 (54.8)
	No Disability	76.2 (43.6)	85.7 (35.9)
Lick	Disability	83.3 (40.8)	66.7 (51.6)
	No Disability	52.4 (51.2)	47.6 (51.2)
Cry	Disability	100.0 (0.0)	50.0 (54.8)
	No Disability	71.4 (46.3)	90.5 (30.1)
Dry	Disability	83.3 (40.8)	83.3 (40.8)

	No Disability	81.0 (40.2)	81.0 (40.2)
Pour	Disability	66.7 (51.6)	66.7 (51.6)
	No Disability	76.2 (43.6)	71.4 (46.3)
Catch	Disability	50.0 (54.8)	50.0 (54.8)
	No Disability	33.3 (48.3)	42.9 (50.7)
Spill	Disability	33.3 (51.6)	66.7 (54.8)
	No Disability	57.1 (50.7)	66.7 (48.3)
Clap	Disability	100.0 (0.0)	100.0 (0.0)
	No Disability	95.2 (21.8)	100.0 (0.0)
Wipe	Disability	66.7 (51.6)	66.7 (54.8)
	No Disability	71.4 (46.3)	61.9 (49.8)
Overall *M*(SD)	Disability	75.9 (21.3)	70.37 (18.1)
	No Disability	69.9 (24.2)	70.4 (21.2)

For children with typical development, a dependent *t*-test found non-significant response latency differences ($t(20) = 1.20$, $p = 0.24$, Cohen's $d = .26$) between animated symbols with environmental sounds ($M = 6169.98$ ms, $SD = 2415.73$) and animated graphic symbols without sound ($M = 5722.74$ ms, $SD = 1489.36$). For children with complex communication needs, a dependent *t*-test found non-significant response latency differences ($t(5) = -0.86$., $p = 0.43$, Cohen's $d = -.35$) between animated symbols with environmental sounds ($M = 6932.21$ ms, $SD = 2063.93$) and animated graphic symbols without sound ($M = 8847.88$ ms, $SD = 6690.61$).

Table 3 provides the average response time for each symbol in each condition for all participants. The biggest difference in response time for participants with typical development was with the symbol *SPLASH*, with faster responses in the no sound condition. The biggest differences in response time for participants with complex communication needs was with the symbol *CATCH*, with responses being faster in the sound condition.

Table 3

Response Latency in ms for all Participants

Symbol	Participant	Sound Response Latency ms(SD)	No Sound Response Latency ms(SD)
Kiss	Disability	7131 (2569)	6369 (2411)
	No Disability	7766 (4802)	5842 (3345)
Splash	Disability	6150 (1876)	7363 (2207)
	No Disability	7478 (3190)	4887 (1079)
Draw	Disability	6028.80 (2143)	6583 (3458)
	No Disability	5784 (4892)	4386 (1932)
Open	Disability	6015 (1860)	7621 (4114)
	No Disability	5834 (3525)	5551 (2814)
Close	Disability	7551 (4924)	6646 (4510)
	No Disability	5395 (1092)	6744 (5841)
Swim	Disability	5107 (2092)	4308 (2650)
	No Disability	3964 (2083)	3815 (1413)
Shake	Disability	6559 (2734)	5595 (1336)
	No Disability	7782 (6219)	6207 (3165)
Sweep	Disability	4029 (2338)	5939 (3386)
	No Disability	4034 (2400)	3766 (1392)

Hit	Disability	14743 (8612)	10102 (0)
	No Disability	6891 (3043)	6621 (2956)
Break	Disability	7736 (5201)	7696 (3130)
	No Disability	5644 (2307)	5539 (1193)
Lick	Disability	8154 (4229)	6205 (3245)
	No Disability	6661 (3963)	4943 (903)
Cry	Disability	7533 (1870)	10407 (3486)
	No Disability	6959 (2354)	5954 (1340)
Dry	Disability	9328 (4153)	9232 (1555)
	No Disability	7483 (2594)	7340 (2183)
Pour	Disability	7415 (5163)	8753 (3720)
	No Disability	7132 (3245)	6976 (2383)
Catch	Disability	8539 (3649)	30040 (48379)
	No Disability	5296 (1837)	5104 (1201)
Spill	Disability	8594 (4419)	27048 (39924)
	No Disability	7009 (2807)	7331 (3697)
Clap	Disability	5980 (3161)	6107 (1992)
	No Disability	5584 (2014)	5842 (2385)
Wipe	Disability	6139 (2466)	6558 (3372)
	No Disability	8581 (4867)	6870 (1782)
Overall *M*(SD)	Disability	6932 (2064)	8848 (6691)
	No Disability	6253 (2591)	5740 (1488)

Symbol Performance Heuristics

Symbol performance heuristics from Schlosser et al. (2011) were applied to this study's symbols. For the children with typical development, 14 of symbols in the sound condition were named (including synonyms) with $\geq$ 70% accuracy and considered satisfactory. Similarly, 12 of symbols in the no sound condition were satisfactory. The second heuristic, name agreement, required participants to name the exact verb (i.e., no synonyms) with $\geq$ 60% accuracy. For the children with typical development, 9 of symbols in both the sound and no sound condition underperformed. The final heuristic, response efficiency (i.e., time to correctly name), was a newly developed criteria. Children with typical development average approximately 6300 ms in their responses in the sound condition and 5700 ms in the no sound condition. Consensus was reached to set this criteria at 6000 ms. For the children with typical development, 8 of symbols in the sound condition were named faster than 6000 ms while 11 of symbols were named in under 6000 ms in the no sound condition. Finally, overall symbol performance was rated as exceptional (5 symbols in the sound condition and 7 symbols in the no sound condition met all three criteria), effective (6 symbols in the sound condition and 3symbols in the no sound condition met two of the three criteria), adequate (4 symbols in the sound conditions and 5 of the symbols in the no sound condition met one of the three criteria), or inadequate (3 symbols in both conditions did not meet any of the criteria). Similar findings were found for children with disabilities, and the results can be found in Table 4 for all participants.

Table 4

Symbol Performance Based on Heuristic Criteria

Symbol	Participant	Exceptional[a]	Effective	Adequate	Inadequate
Kiss	Disability		Both conditions[b]		
	No Disability	No sound[c]	Sound		
Splash	Disability				Both conditions
	No Disability			No sound	Sound
Draw	Disability		Both conditions		
	No Disability		Both conditions		
Open	Disability		Both conditions		
	No Disability	Both conditions			
Close	Disability		Both conditions		
	No Disability		Sound[d]	No sound	
Swim	Disability	Both conditions			
	No Disability	Both conditions			
Shake	Disability			Both conditions	
	No Disability		Both conditions		
Sweep	Disability	Both conditions			
	No Disability	Both conditions			

Hit	Disability	Both conditions
	No Disability	Sound ← No sound
Break	Disability	Both conditions
	No Disability	Both conditions
Lick	Disability	Sound [e] ← No sound
	No Disability	No sound ← Sound
Cry	Disability	Sound ← No sound
	No Disability	No sound ← Sound
Dry	Disability	Both conditions
	No Disability	Both conditions
Pour	Disability	Both conditions
	No Disability	Both conditions
Catch	Disability	Both conditions
	No Disability	Both conditions
Spill	Disability	No sound ← Sound
	No Disability	Both conditions
Clap	Disability	Sound ← No sound
	No Disability	Both conditions
Wipe	Disability	Both conditions

Sound ← No sound

Note. a = heuristics: naming accuracy $\geq$ = 70%, name agreement $\geq$ = 60%, and response time $\leq$ 6000 ms; b = red ovals indicate no heuristic performance difference between the symbol conditions; c = black ovals indicate the symbol without sound performed better than the symbol with sound by one heuristic category; d = blue ovals indicate the symbol with sound performed better than the symbol without sound by one heuristic category; e = green ovals indicate a two-category improvement where symbols with sound performed better than symbols without sound.

Symbol Error Analysis

The error analysis is summarized in Table 5. The percentages of total errors made in each condition are listed for each error category and type. The error analysis for typically developing participants revealed that sematic type errors were the most common in both conditions. Superordinate errors were most common semantic type error in the no sound condition and the coordinate errors were the most common in the sound condition. The findings for participants with complex communication needs were classified most frequently as “other type errors” in both conditions. However, sematic type errors were made at a higher rate in the sound condition than in the no sound condition.

Table 5

Symbol Error Analysis Patterns for all Participants, Expressed as a Percentage of Total Errors

		Percentage	
Error Categories and types	Participant	Sound	No sound
Semantic type errors			
Semantic type coordinate errors	Disability	13.51	17.07
	No Disability	18.83	15.33

Semantic type superordinate errors	Disability	13.51	9.76
	No Disability	13.64	22.67
Semantic type associative errors	Disability	18.92	7.32
	No Disability	16.88	16.67
Semantic type errors total	Disability	45.95	34.15
	No Disability	49.35	54.67
Visual type errors			
Frank visual errors	Disability	2.70	12.20
	No Disability	1.30	5.33
Misinterpretation of picture errors	Disability	0.00	0.00
	No Disability	1.30	2.00
Visual type errors total	Disability	2.70	12.20
	No Disability	2.60	7.33
Auditory type errors			
Frank auditory errors	Disability	2.70	N/A
	No Disability	0.00	N/A
Misinterpretation of sound errors	Disability	0.00	N/A
	No Disability	0.00	N/A
Auditory type errors total	Disability	2.70	N/A
	No Disability	0.00	N/A
Other errors			
Verb to adjective errors	Disability	0.00	2.44
	No Disability	9.74	7.33

Verb to noun errors	Disability	0.00	0.00
	No Disability	1.30	0.00
Mixed errors	Disability	0.00	4.88
	No Disability	25.32	20.00
Auditory and visual	Disability	0.00	N/A
	No Disability	0.00	N/A
Unrelated errors	Disability	45.95	46.34
	No Disability	7.79	8.67
Omission errors	Disability	2.70	0.00
	No Disability	3.90	2.00
Other errors total	Disability	48.65	53.66
	No Disability	48.05	38.00

Chapter 4: Discussion

A comparison between this study and the results from Harmon et al. (2014) is warranted because the same procedures, stimuli, and dependent variables were used. A comparison of the results may reveal novel differences between 3- and 4-year-olds as well help us better understand the role of animation and sound as children age. It is important to note that this comparison is somewhat haphazard because each study took place in different locations with very different populations. Additionally, the standardized assessments administered were very different. Therefore, the comparison must be interpreted with caution.

Naming and Response Latency

The results of this study are in direct contrast with previous results from Harmon et al. (2014). Specifically, Harmon et al. found that 3-year-old children had significantly greater verb naming accuracy in the sound condition when compared to the no sound condition. In the current study, 4-year-old verb naming accuracy was essentially the same in each condition. The discrepancy between the results with 3-year-old and 4-year-old participants has many potential explanations.

First, the stimuli used in each study were the same in every way, but it is possible that embedded sounds are not the mediating factor leading to greater naming accuracy by our 4-year-old participants. Rather, it is possible that animation is the most important factor to enhance the iconicity of a symbol. Neither this study nor Harmon et al. (2014) compared static symbols and animated symbols with embedded sound, limiting our ability to isolate these contextualized effects.

Second, children may become less reliant on the sounds for contextual support as indicated by the higher percent naming accuracy of the 4-year-olds when compared to 3-year-olds from Harmon et al. (2014) in the no sound condition. Interestingly, 3-year-olds (M = 73.91) named verbs in the sound condition at similar rates as 4-year-olds (M = 74.60) in the current study, possibly negating the validity of an age effect. Although, it is still possible that the current participants did have stronger language skills resulting in an age effect. The 4-year-old children may have had a more detailed understanding of each target word because as language continues to develop, the semantic network becomes more robust. As a result, if the symbol did not represent their understanding of the target word they may have labeled it with a different word (Bowerman, 1978). In contrast, the 3-year-old children may have had a broader understanding of each target verb thus allowing them to give it as response when presented with a symbol that had any notable relationship with the target verb (Reich, 1976). To substantiate this claim, a discussion of the differences between fast mapping and slow mapping is required.

Fast mapping occurs when a child assigns meaning to a word based upon his first encounter with it, resulting in a broad definition (Carey & Bartlett, 1978). Over time, the children are exposed to the word several more times during which slow mapping occurs. Slow mapping allows the child to refine and deepen his understanding of the word (Swingley, 2010). The fact that the 4-year-old children have had an additional year of exposure to language suggests that they have a more complete understanding of the target verbs, which may explain the fact that the sound was not as helpful to this group as it was to the 3-year-old children.

Alternatively, the differences in the results could be explained by the intersection of sociocultural aspects, context, and symbol iconicity. Harmon et al. collected data in the greater Boston area, and Massachusetts consistently ranks higher than Idaho in educational outcomes for Pre-K through 12th grade (Ziegler, 2021). Additionally, many of the verbs used in this study require context for accurate naming, that is, verbs such as dry, shake, and lick are associated with a specific scenario or previous experience. The symbol for dry included water on a surface and a character with a towel; however, to many children, that is associated with cleaning. Similarly, the symbol for lick included a lollipop, but it is unclear how strong of a relationship lick and lollypop have. Perhaps an animal licking the character's face would have a stronger relationship between the word and its referent. Overall, sociocultural factors, symbol context, and iconicity cannot be ignored when selecting appropriate symbols.

Finally, the amorphous blob, which changes its shape to portray various nouns (e.g., swimming pool), may have led to some confusion. The blob was incorporated into the ALP set because the creators wanted the verb with its associated movement to be the most important part. Additionally, the blob was to take on familiar contexts, but not be so specific to any one context. To support this rationale, all 4-year-olds comprehended the verbs in this study with 100% accuracy as indicated by the screening task. The difference between the screening task and the experimental task was context. That is, the experimenters utilized toy props in a play format to elicit labelling or receptive knowledge of the verbs while the experimental task utilized the blob for context. Additionally, children in this study were frequently confused making comments such as "Why is he eating garbage/rocks?" or "Why is he living in a cloud?" While the rationale for the blob is sound, language is contextually-based, and future symbol sets and symbol selection for AAC systems must account for this context as well as individual client characteristics (e.g., culture).

The participants with developmental disabilities had a higher rate of naming accuracy in the sound condition (i.e., 5.56% greater accuracy), but not to a statistically significant degree. Additionally, Cohen's *d* indicated that the difference in accuracy resulted in a large effect size. However, no causal inferences can be made at this time about the impact of sound on the accurate naming of animated symbols by children with developmental disabilities due to the small sample size in this study. Regardless, it appears this line of inquiry may be valuable to those who require AAC.

While not significant, the response latency for the 4-year-old participants was marginally longer in the sound condition when compared to the no sound condition resulting in a small effect size (i.e., d = .26). The slower responses in the sound condition may be due to the time required to process the additional environmental sound stimuli; however, that processing time appears negligible at less 700ms. Response latency results suggest that the inclusion of sound did not overburden the cognitive systems involved in processing and subsequently naming the verbs. That is, the visual motion perception system as well as auditory stimulus interpretation regions of the brain were likely capable of handling the dual modality input for 4-year-old participants as well as the children with disabilities. It is also important to note that the sound embedded in each symbol played 2s into the animation. This slight delay may have been responsible for the slower response latency in the sound condition. This delay was not altered for experimental purposes because it would have shortened the duration of animated symbol with embedded sound, making it different from the no sound symbol condition. However, the researcher did observe children occasionally naming symbols in the sound condition before the sound started. In these instances, the child was not exposed to the sound before forming a response. This suggests that 4-year-old children may be more impactful than the actual sound.

Symbol Error Analysis

Similar to Harmon et al. (2014), semantic type errors and other errors were the most common in each condition for the 4-year-old participants with typical development. Interestingly, the addition of sound appears to have decreased the number of superordinate errors (e.g., swimming for splashing) made 4-year-olds by almost 10%. Similarly, coordinate errors (i.e., approved synonyms) were slightly lower in the sound condition. These results suggest that symbols paired with animation and sound could (a) reduce the number of children that need instruction for a specific verb and (b) decrease the time required to teach difficult to learn verbs. However, this trend was not true for the small sample of participants with a disability. In fact, this group made fewer semantic errors in the no sound condition suggesting that sound might not have the same facilitative effects it did for the 4-year-old participants with typical development. The pilot data for the children with disabilities is too small to make any conclusive statements. With respect to the visual errors, all participants regardless of disability made fewer errors in the sound condition than in the no sound condition. As Harmon et al. (2014) noted, the sound highlighted the most important aspect of the animation, allowing the participants to focus on the action rather than extraneous content within the symbol. Finally, the other errors category reflected minor differences between the two conditions. In contrast to Harmon et al. (2014), 4-year-old children generally had fewer errors classified as "other" in the sound condition. For example, the verb-to-noun and verb-to-adjective errors were greater in the sound condition than in the no sound condition. While Harmon et al. (2014) suggested that sound could assist children with differentiating word class, the current data do not support that hypothesis for 4-year-olds. Perhaps these results are evidence of an age effect, indicating that as the language of children with typical development further develops, they are less reliant on environmental sounds to label the symbol. Alternatively, it is possible that the results are different because the current sample size is smaller.

Clinical Implications

While children with typical development are not individuals who would use AAC symbols, this study provides data on the general effectiveness of environmental sound as a teaching tool. Specifically, Harmon and colleagues (2014) found that naming accuracy of verbs was better for animated symbols embedded with environmental sound, while this study found so such effect. This suggests several possible explanations such as sociocultural factors (e.g., educational outcome differences between the two samples), poor symbol iconicity, or an age effect. Therefore, it is imperative that clinicians account for these variables when selecting symbols or symbol sets to incorporate into any AAC system. Context is a powerful tool in any symbol, and the amorphous blob utilized in the ALP set sometimes has strong context promoting the symbol-referent relationship (e.g., kiss); however, in other instances (e.g., dry), the context is unclear and the symbol-referent relationship is open for interpretation.

Next, participants' performance varied widely in each condition as indicated by the standard deviations. This is important because it suggests that symbols incorporating animation or animation with embedded sound may be beneficial for certain populations. Therefore, clinicians must carefully evaluate a client's current level of language function, cognitive function, and learning preferences when deciding if graphic symbols incorporating animation and sound would be facilitative in learning new words or constructs.

The results from this study and Harmon et al. (2014) and generally supportive of animation and sound; however, it would be unreasonable to have a display set of looping animations and sounds. Rather, these tools would likely be appropriate as a just-in-time prompt (Schlosser et al., 2016) to support learning. That is, the animation or sound would only activate if the client or the SLP "swiped" across the symbol to start those features. Eventually this just-in-time prompt would fade as the client increases their understanding of the symbol. Finally, both studies exclusively focused on verbs in which participants had demonstrated understanding. It is possible that environmental sound would be beneficial when introducing a symbol for which the child does not have in the semantic repertoire.

Limitations

While the current study did not yield statistically significant differences for the 4-year-old participants, the results are not conclusive for children with disabilities. First, these children with developmental disabilities represent pilot data and had a variety of different diagnoses (ASD, Down syndrome, and intellectual impairment) as well as a wide age range (5 to 16 years old). The results for children with disabilities are likely skewed by the older and/or higher functioning children included. Additionally, this study included an extremely small sample size of children with disabilities and a relatively small sample of 4-year-old participants. A power analysis revealed that 27 4-year-olds would be required for a medium effect size (Cohen's $d = .5$) with an alpha level of .05. However, it is unlikely that the additional children with typical development would change the verb naming results or response latencies given the negligible mean difference between the sound and no sound conditions.

Finally, as Harmon et al. (2014) discussed, there is no standard for environmental sounds in AAC. However, Gygi and Shafiro (2010) introduced the Database for Environmental Sound Research and Application (DERSA) as a way to help researchers standardize sounds for experimental purposes. Perhaps future studies can use those sounds, but the purpose of this study was to extend the results from Harmon et al. (2014) to 4-year-old children. Therefore, both studies used the same sounds and animations. Thus, there is no way to determine if the sounds used in this study accurately represented the action, making it possible that different sounds may lead to different results.

Future Directions

Future research should investigate the impact of sound with other AAC symbols sets, sounds from DERSA, and children with disabilities. Additionally, animation and sound are proving to be highly useful for both children with typical development (Harmon et al., 2014; Schlosser et al., 2014) and those with developmental disabilities (Fujisawa et al., 2011; Schlosser et al., 2019). While this study cannot conclude that sound is better than no sound, the individual participant performances varied greatly in each condition. This suggests that the usefulness of animated graphic symbols with embedded sound may be client dependent. Therefore, researchers should consider individual level demographics, cognitive ability, and language performance more closely in future work.

A final avenue for future research is investigating the impact of adding sound to graphic symbols on a child's ability to learn a new verb. The current study and Harmon et al. (2014) exclusively focused on the verbs the children already understood. Moving forward, it is important to understand if adding sound to symbols makes it easier for those with language impairments to learn new verbs. This direction of research would be of particular clinical importance as it has been established that verbs need to be a higher priority in AAC therapy (Schlosser et al., 2019).

Conclusion

In sum, animated symbols with embedded sounds were not named more readily or faster when compared to the no sound condition. However, there was variability in the participants' performance in each condition. Clinicians should consider client-level variables (e.g., semantic knowledge and cognition) prior to implementing animation or sound. With that said, animated symbols with sound was not detrimental to naming accuracy and may be used clinically with caution. Finally, it is unclear if animation and sound could enhance the learning of new words given the contextual clues they provide. Professionals may want to consider how sound and animation affect learning.

References

Angermeier, K., Schlosser, R. W., Luiselli, J. K., Harrington, C., & Carter, B. (2008). Effects of iconicity on requesting with the Picture Exchange Communication System in children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 2*(3), 430-446. doi:10.1016/j.rasd.2007.09.004

Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education, 101*, 150-167. doi:10.1016/j.compedu.2016.06.005

Bertelson, P., & De Gelder, B. (2004). Crossmodal space and crossmodal attention. In 1220921776 908418863 C. Spence & 1220921777 908418863 J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 141-178). Oxford: Oxford University Press.

Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: A review. *Le Travail Humain: A Bilingual and Multi-Disciplinary Journal in Human Factors, 63*(4), 311-329.

Bloomberg, K., Karlan, G. R., & Lloyd, L. L. (1990). The comparative translucency of initial lexical items represented in five graphic symbol systems and sets. *Journal of Speech, Language, and Hearing Research, 33*(4), 717-725. doi:10.1044/jshr.3304.717

Bowerman, M. (1978). Systematizing semantic Knowledge: Changes over time in the Child's organization of word meaning. *Child Development, 49*(4), 977-988. doi:10.2307/1128737

Braddick, O., Birtles, D., Wattam-Bell, J., & Atkinson, J. (2005). Motion- and orientation specific cortical responses in infancy. *Vision Research, 45*(25-26), 3169-3179. doi:10.1016/j.visres.2005.07.021

Cardon, T., & Azuma, T. (2012). Visual attending preferences in children with autism spectrum disorders: A comparison between live and video presentation modes. *Research in Autism Spectrum Disorders, 6*(3), 1061-1067. doi:10.1016/j.rasd.2012.01.007

Chen, Y., & Spence, C. (2010). When hearing the bark helps to identify the dog: Semantically congruent sounds modulate the identification of masked pictures. *Cognition, 114*(3), 389-404. doi:10.1016/j.cognition.2009.10.012

Driver, J., & Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. *Current Biology, 10*(20). doi:10.1016/s0960-9822(00)00740-5

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., . . . Reilly, J. S. (1994). *MacArthur-Bates Communication Development Inventory*. Baltimore: Paul H. Brookes Publishing.

Fujisawa, K., Inoue, T., Yamana, Y., & Hayashi, H. (2011). The effect of animation on learning action symbols by individuals with intellectual disabilities. *Augmentative and Alternative Communication, 27*(1), 53-60. doi:10.3109/07434618.2011.553245

Gaver, W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction, 2*(2), 167-177. doi:10.1207/s15327051hci0202_3

Gygi, B., & Shafiro, V. (2010). Development of the database for environmental sound research and Application (desra): Design, functionality, and Retrieval considerations. *EURASIP Journal on Audio, Speech, and Music Processing, 2010*(1), 654914. doi:10.1186/1687-4722-2010-654914

Harmon, A. C., Schlosser, R. W., Gygi, B., Shane, H. C., Kong, Y., Book, L., . . . Hearn, E. (2014). Effects of environmental sounds on the guessability of animated graphic symbols. *Augmentative and Alternative Communication, 30*(4), 298-313. doi:10.3109/07434618.2014.966206

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*(6), 722-738. doi:10.1016/j.learninstruc.2007.09.013

Kenney, M. K., & Kogan, M. D. (2011). Special needs children with speech and hearing difficulties: prevalence and unmet needs. *Academic pediatrics*, *11*(2), 152-160. doi.org/10.1016/j.acap.2011.01.003

Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the Colavita visual dominance effect. *Experimental Brain Research, 184*(4), 533-546. doi:10.1007/s00221-007-1120-z

Lee, H., & Hong, K. (2013). The effect of animations on the AAC symbol recognition of children with intellectual disabilities. *Special Education Research, 12*(2), 185. doi:10.18541/ser.2013.06.12.2.185

Masterson, J., Druks, J., & Gallienne, D. (2008). Object and action picture naming in three- and five-year-old children. *Journal of Child Language, 35*(2), 373-402. doi:10.1017/s0305000907008549

Mineo, B. A., Peischl, D., & Pennington, C. (2008). Moving targets: The effect of animation on identification of action word representations. *Augmentative and Alternative Communication*, *24*(2), 162-173. doi:10.1080/07434610802109915

Nava, E., Grassi, M., & Turati, C. (2016). Audio-visual, visuo-tactile and audio-tactile correspondences in preschoolers. *Multisensory Research*, *29*(1-3), 93-111. doi:10.1163/22134808-00002493

Ogletree, B. T., Mcmurry, S., Schmidt, M., & Evans, K. (2018). The changing world of augmentative and alternative communication (AAC): Examining three realities faced by todays AAC provider. *Perspectives of the ASHA Special Interest Groups,3*(12), 113-122. doi:10.1044/persp3.sig12.113

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Ploetzner, R., & Lowe, R. (2012). A systematic characterisation of expository animations. *Computers in Human Behavior*, *28*(3), 781-794. doi:10.1016/j.chb.2011.12.001

Reich, P. (1976). The early acquisition of word meaning. *Journal of Child Language, 3*(1), 117-123. doi:https://doi.org/10.1017/S0305000900001343

Schlosser, R. W., Shane, H., Sorce, J., Koul, R., & Bloomfield, E. (2011). Identifying performing and underperforming graphic symbols for verbs and prepositions in animated and static formats: A research note. *Augmentative and Alternative Communication, 27*(3), 205-214. doi:10.3109/07434618.2011.592217

Schlosser, R. W., Shane, H., Sorce, J., Koul, R., Bloomfield, E., Debrowski, L., . . . Neff, A. (2012). Animation of graphic symbols representing verbs and prepositions: Effects on transparency, name agreement, and identification. *Journal of Speech, Language, and Hearing Research, 55*(2), 342-358. doi:10.1044/1092-4388(2011/10-0164)

Schlosser, R. W., Koul, R., Shane, H., Sorce, J., Brock, K., Harmon, A., . . . Hearn, E. (2014). Effects of Animation on Naming and Identification Across Two Graphic Symbol Sets Representing Verbs and Prepositions. *Journal of Speech, Language, and Hearing Research, 57*(5), 1779-1791. doi:10.1044/2014_jslhr-l-13-0193

Schlosser, R. W., Shane, H. C., Allen, A. A., Abramson, J., Laubscher, E., & Dimery, K. (2016). Just-in-time supports in augmentative and alternative communication. *Journal of Developmental and Physical Disabilities, 28*(1), 177-193.

Schlosser, R. W., Brock, K., Koul, R.K., & Shane, H. C. (2019). Does animation facilitate understanding of graphic symbols representing verbs in children with autism spectrum disorder? *Journal of Speech, Language, and Hearing Research.*

Schopler, E., Van Bourgondien, M., Wellman, G., & Love, S. (2010). *Childhood Autism Rating Scale, Second Edition*. Torrance, CA: WPS

Sparrow, S. S., Cicchetti, D.V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales, Third Edition*. San Antonio, TX: Pearson

Sussman, E. (2016). Auditory scene analysis: An attention perspective. *The Clinical Research Education Library, 1*(1), 2989-3000. doi:10.1044/cred-pvd-c16002

Sussman, E. S. (2017). Auditory scene analysis: An attention perspective. *Journal of Speech, Language, and Hearing Research, 60*(10), 2989-3000. doi:10.1044/2017_JSLHR-H-170041

Swingley, D. (2010). Fast mapping and slow mapping in children's word learning. *Language Learning and Development, 6*(3), 179-183. doi:10.1080/15475441.2010.484412

Zeigler, B. (2021). Education Rankings: Measuring how well states are educating their students. *US News.* https://www.usnews.com/news/best-states/rankings/education

Appendix A

List and Definition of Error Categories and Types

Semantic-type errors

- Coordinate errors: Response has semantically and frequently visual similarities to target and/or auditory similarities.
- Superordinate errors: Response demonstrates an overgeneralization.
- Associative errors: Response has a contextual relationship with the target.

Visual type errors

- Frank visual errors: Response is free from any semantic connection to the target, but something about the graphic prompted the response.
- Misinterpretation of symbol errors: Response names an action pictured in the graphic that is not the target.

Auditory-type errors

- Frank auditory errors: Response is free from any semantic connection to the target, but something about the sound prompted the response.
- Misinterpretation of sound errors: Response identified an aspect of the sound that is not the target.

Other errors

- Verb-to-noun errors: Response labels the object involved rather than the action.
- Verb-to-adjective errors: Response consists of a related adjective rather than a verb.
- Mixed errors: Response fit into two or more error categories.
- Auditory and visual errors: The response contained both an auditory and a visual error.
- Unrelated errors: A response that does not appear to have any connection to the target.

- Omission errors: Lack of response or a response claiming lack of knowledge.

Appendix B

Procedures:

1. For procedural reliability, place a ✓(completed) or X (did not complete). For child response place a **+** for a correct response or a **-** for an incorrect response. For responses that are wrong, **write in response**.
2. Ensure Zoom is set to “Gallery view” so both participants in the Zoom call are seen side-by-side.
3. Experimenter begins screen recording.
4. Select share screen, with sound, and PPT.
5. Verify full screen and they can see your screen.
6. Say, “Minimize our videos and move the box out of the way.”
7. Ask the parent to ensure their phone is on airplane mode.
8. Tell the parent to start filming their child and the screen.
9. Proceed to familiarization task slide.
10. Following the first familiarization trial, confirm with both the parent and child that they can hear you and the PPT.
11. Affirm/negate responses using script on slide.
12. Repeat familiarization trials until both are correct.