

Article

Not peer-reviewed version

QuickPic AAC: An AI-Based Application to Enable Just-in-Time Generation of Topic-Specific Displays for Persons Who Are Minimally Speaking

[Christina Yu](#)*, [Ralf Schlosser](#), Maurício Fontana de Vargas, [Leigh Anne White](#), [Rajinder Koul](#), [Howard Shane](#)

Posted Date: 15 July 2024

doi: 10.20944/preprints202407.0999.v1

Keywords: autism spectrum disorder (ASD); artificial intelligence (AI); augmentative and alternative communication (AAC); speech pathology; application; topic displays; just-in-time (JIT)



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

QuickPic AAC: An AI-Based Application to Enable Just-in-Time Generation of Topic-Specific Displays for Persons Who Are Minimally Speaking

Christina Yu ^{1,2,*}, R. W. Schlosser ^{1,3}, Maurício Fontana de Vargas ⁴, Leigh Anne White ¹, Rajinder Koul ⁵ and Howard C. Shane ^{1,2}

¹Boston Children's Hospital, Waltham MA

²Massachusetts General Hospital Institute of Health Professionals, Boston MA;

³Northeastern University, Boston MA;

⁴McGill University, Quebec, Canada

⁵University of Texas at Austin, Austin TX

*Correspondence: christina.yu@childrens.harvard.edu; Tel.: 781-216-1075

†This is a draft to be submitted to the International Journal of Environmental Research and Public Health.

Abstract: As artificial intelligence (AI) makes significant headway in various arenas, the field of Speech-Language Pathology is at the precipice of experiencing a transformative shift towards automation. This study introduces *QuickPic AAC*, an AI-driven application designed to generate topic-specific displays from photographs in a "just-in-time" manner. Using *QuickPic AAC*, this study aimed to (a) determine which of two AI algorithms (NLG-AAC and GPT-3.5) results in greater specificity of vocabulary (i.e., percentage of vocabulary kept/deleted by clinician relative to vocabulary generated by *QuickPic AAC*; percentage of vocabulary modified); and to (b) evaluate perceived usability of *QuickPic AAC* among practicing speech-language pathologists. Results revealed that the GPT-3.5 algorithm consistently resulted in greater specificity of vocabulary and that speech-language pathologists expressed high user satisfaction for the *QuickPic AAC* application. These results support continued study of the implementation of *QuickPic AAC* in clinical practice and demonstrate the possibility of utilizing topic-specific displays as just-in-time supports. *QuickPic AAC: An AI-Based Application to Enable Just-in-Time Generation of Topic-Specific Displays for Persons Who Are Minimally Speaking.*

Keywords: autism spectrum disorder (ASD); artificial intelligence (AI); augmentative and alternative communication (AAC); speech pathology; application; topic displays; just-in-time (JIT)

Introduction

With the advent of mobile technology, the use of applications ("Apps") in augmentative and alternative communication (AAC) has become integral to the standard of care for persons who are minimally speaking (Shane et al., 2012) (individuals who are minimally-speaking may include persons with developmental disabilities [e.g., autism, intellectual disabilities, etc.], acquired disorders [e.g., aphasia, Traumatic Brain Injury], progressive disorders [e.g., Muscular Dystrophy], and temporary conditions [e.g., recovering from surgery in the Intensive Care Unit] (Beukelman & Light, 2020). Many apps provide a range of tools that serve as a communication platform as well as a medium to provide language support. *QuickPic AAC* is a new and innovative app that seamlessly blends artificial intelligence (AI) and visual supports to empower minimally-speaking individuals who require support generating utterances by selecting graphic representations or text from a display.

QuickPic AAC harnesses the power of AI to interpret visual scenes from a photograph, allowing it to identify characters and their actions. The source of the picture scenes can come from a photo library, a fresh photo snapshot, or an internet search. *QuickPic AAC* then transforms the visual input into a mixed display which is a combination of the visual scene (photo) and the elements of vocabulary thematically related to the scene arranged in a grid display (Shane et al., 2014). The grid

display is arranged in the form of a modified Fitzgerald Key that parses and color-codes the grammatical parts of a sentence (Thistle & Wilkinson, 2009). *QuickPic AAC* has the following categories from left to right: pronouns, verbs, prepositions, adjectives, and objects. In other words, after analyzing the photo, *QuickPic AAC* constructs a grid that strategically places symbols representing the subjects and their activities in the scene. Notably, the app uses facial recognition to recognize individuals and retains this knowledge to accurately identify individuals in future mixed displays. *QuickPic AAC* also allows instructors to edit and customize the symbols in the grid, ensuring the most accurate representation of the scene. This collaborative and customizable aspect guarantees that the app's generated vocabulary not only aligns with the visual content, but is also personalized and meaningful to its user enabling learners to better grasp language concepts.

The thematic or topic-specific vocabulary that is arranged in grammatical categories (i.e., based on the Fitzgerald Key) is known as a topic specific display (TSD). TSDs are a type of aided approach that enable users to communicate appropriately with phrase production in the context of a particular activity (Goossens' & Crain, 1986; Goossens' et al., 1992) through the arrangement of linguistic elements on a single page for constructing a sentence.

Traditionally, developing meaningful and functional TSDs has required significant advanced planning and programming, and therefore, time, from mentors working with individuals who are minimally-speaking. For example, imagine a teacher is planning to introduce a new science lesson on a particular forest biotope in the weeks ahead. In addition to planning the lesson in general terms, this teacher would need to gather all the vocabulary needed for the student who is minimally-speaking so that the student can be an effective participant in that lesson, and then organize it in an intuitive way.

Because the *QuickPic AAC* app enables automatic generation and organization of vocabulary from a single photograph, one could upload a photo of a forest biotope and the app would automatically generate and organize the vocabulary in the form of graphic symbols (e.g., Picture Communication Symbols). If functional, the mentor would save considerable time and elevate TSDs into the realm of just-in-time supports (JITs) (O'Brien et al., 2016; O'Brien et al., 2017; Schlosser et al., 2016; 2017). This is something that was previously unthinkable given the advanced planning and time-consuming preparations.

Usability testing is a critical element in the product development cycle of apps in mobile health (mHealth) and education (Zapata et al., 2015). There are a host of methods available for usability testing, including questionnaires, think aloud walkthrough, task completion, interviews, focus groups, heuristic testing, and automated methods. A recent scoping review (Maramba et al., 2019) revealed that most usability studies in eHealth use a combination of at least two of these methods and this was the overall order in terms of frequency of use: questionnaires (n=105), task completion (n=57), 'Think-Aloud' (n=45), interviews (n=37), heuristic testing (n=18) and focus groups (n=13).

Using a combination of quantitative and thematic analyses methods, the purpose of this study was to (a) determine which of two AI algorithms (NLG-AAC and GPT-3.5) results in more relevant vocabulary with the *QuickPic AAC* application; and to (b) evaluate the perceived usability of *QuickPic AAC* among practicing speech-language pathologists.

Methods

Participants

Participants included eight speech-language pathologists (SLPs) ranging in age range from 25 to 64 years based in an outpatient pediatric hospital: Four participants were between 25 and 34 years old, three participants were between 35 and 44 years old, and one participant was between 55 and 64 years old. In order to be included, participants had to meet the following criteria: (a) an active American Speech-Language-Hearing Association (ASHA) Certificate of Clinical Competence for Speech-Language Pathologists (CCC-SLP); (b) a minimum of one year of experience working with individuals who use AAC or individuals who might benefit from AAC; and (c) experience in having created at least one TSD. Participants were recruited based upon a convenience sampling within an

outpatient AAC center in the Northeast of the United States. Table 1 provides an overview of participant characteristics.

Table 1. Participant characteristics.

Participant	Race	Ethnicity	CA	Years Practicing as an SLP	Frequency of Working with Individuals Who Use AAC	Have you created topic displays	Frequency of creating topic displays	Average length to create topic displays
1	White	Not Hispanic or Latino	25-34	2	Weekly	Yes	Occasionally	31-40 minutes
2	More than one race	Not Hispanic or Latino	25-34	4	Daily	Yes	Weekly	11-20 minutes
3	White	Not Hispanic or Latino	35-44	17	Daily	Yes	Monthly	<10 minutes
4	White	Not Hispanic or Latino	35-44	12	Daily	Yes	Weekly	11-20 minutes
5	White	Not Hispanic or Latino	25-34	2	Weekly	Yes	Monthly	21-30 minutes
6	White	Not Hispanic or Latino	35-44	12	Daily	Yes	Daily	<10 minutes
7	White	Not Hispanic or Latino	25-34	6	Daily	Yes	Monthly	11-20 minutes
8	White	Not Hispanic or Latino	55-64	35	Monthly (varies)	Yes	Occasionally	51-60 minutes

Note. CA = chronological age; SLP = speech language pathologist; AAC = augmentative and alternative communication.

The Institutional Review Board considered this study as exempt because it is limited to research activities in which the disclosure of the human subjects' responses outside the research did not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation. Participants provided verbal consent.

Materials

Materials included (a) a tablet (i.e., iPad Pro) and the *QuickPic* AAC iOS application; (b) *QuickPic* AAC Reference Guide (see Appendix A); (c) the Demographic and AAC Experience Questionnaire; (d) a vignette; (e) photographs; and (f) two usability questionnaires.

Tablet and *QuickPic* AAC. The *QuickPic* AAC app ran on an iPad Pro. The *QuickPic* AAC app evolved from the development of an earlier prototype described in Fontana de Vargas et al. (2022). To generate vocabulary automatically, *QuickPic* AAC employs two different approaches for generating vocabulary. The first approach, proposed by Fontana de Vargas and Moffat (2021) (which has now been coined, NLG-AAC), uses the Visual Storytelling Dataset (VIST) (Huang et al., 2016) as the main source of vocabulary. VIST is composed of 65,394 photos of personal events, grouped in 16,168 stories. Each photo is annotated with captions and narrative phrases that are part of a story, created by Amazon Mechanical Turk workers. The NLG-AAC method works by first identifying the photographs in VIST that are most similar to the input photograph. This is accomplished by calculating the sentence similarity between the input photo caption, generated using the computer vision technique from Fang et al. (2015), and all VIST photos captions. Then, the method retrieves all stories associated with those photographs and finds the most relevant words to present in *QuickPic* AAC by applying the Affinity Propagation clustering algorithm (Frey & Dueck, 2007), and finally, gathering the most frequent words in the identified clusters.

The second approach, named GPT-AAC, takes advantage of recent advancements in Natural Language Processing (NLP), a subfield of AI. More specifically, the method prompts the Large Language Model (LLM) GPT-3.5 to produce the desired set of words related to the input photo caption (which is created using the method from Fang et al. (2015), as in NLG-AAC). The prompt used by the method is shown below:

"You are a Speech Language Pathologist specialized in Augmentative and Alternative Communication."

"Your task is to provide vocabulary related to a situation to help a person with communication disability to formulate messages about the situation. This vocabulary must contain words that people would often use to talk about that situation, either to describe it as well as to tell a story about it."

"The vocabulary must contain 20 verbs, 20 descriptors (adjectives and adverbs not terminating with LY), 20 objects, and 15 prepositions."

"All words must be in the first person singular, infinitive form without 'to'."

The QuickPic AAC Reference Guide. The *QuickPic* AAC Reference Guide is a set of instructions that is available within the app (see Appendix A). For the purposes of this paper, the following terminology is adopted to describe aspects of *QuickPic* AAC communication displays (see Figure 1): (a) Topic Specific Display: thematic or topic-specific vocabulary that is arranged in grammatical categories (subject, verb, object, etc.); (b) Static Scene Cue: a photograph of a single activity and/or concept (Schlosser et al., 2013; Shane et al., 2014); (c) Mixed Display: a display containing a scene cue combined with a topic specific display (Shane et al., 2014).



Figure 1. Guide describing components of a finalized communication display within QuickPic.

Demographic and AAC Experience Questionnaire. The Demographic and AAC Experience Questionnaire elicited key demographic data (e.g., years as a practicing SLP) and previous experience with AAC, including their perspectives on TSDs.

Vignette. The Vignette was a prewritten case study that informed participants of the context in which they would be creating the TSDs. This was provided to all participants to read prior to creation of a TSD with *QuickPic* AAC:

You are a speech-language pathologist in an outpatient pediatric setting and have a 7;2 year old male patient with a primary diagnosis of autism spectrum disorder- level 3. Medical history includes no functional concerns regarding vision, hearing, or motor status. Receptive language skills include strong comprehension of noun-based vocabulary and ability to follow single-step directions within familiar contexts. Expressive language skills include scripted phrases (e.g., I want __), and single word approximations to label. Aided communication strategies include a grid-based communication application used primarily for requesting, labeling, and protesting. A goal of speech therapy is commenting/describing using 3 word utterances. A highly preferred activity/topic of conversation are cars/trains. Based upon this case study, create a QuickPic AAC display revolved around cars/trains using the 'search [the web]' function.

Photographs. As noted in the instructions within the Vignette, participants were asked to choose one photograph for both conditions using the “Search the Web” feature of *QuickPic AAC*. One participant each chose a photo of a sports car on the road, and a photo of two boys playing trains together, respectively. Two participants each chose a photo of two boys playing race cars together, a photo of a boy playing with a wooden train set on the floor, and a photo of a boy playing with cars and trucks on the hardwood floor, respectively. Some participants chose identical photos from the web searches, likely because these images appeared first as initial search results.

Two Usability Questionnaires. The Mobile Health (mHealth) App Usability Questionnaire (MAUQ) (Zhou et al., 2019) and a questionnaire adapted from Fontana de Vargas et al. (2022) were administered. The MAUQ (Zhou et al., 2019; see Appendix B) was used to assess the usability of *QuickPic AAC* with its two approaches. The MAUQ has adequate psychometric characteristics and includes a 7-point Likert scale containing 16 items about interaction, vocabulary and usage factors. The MAUQ was adapted minimally to meet the specific needs of our user study. Specifically, one question was eliminated (i.e., I could use the app even when the Internet connection was poor or not available) as the *QuickPic AAC* application requires Internet connectivity. Additionally, one question was modified from “This mHealth app provides an acceptable way to deliver healthcare services, such as accessing educational materials, tracking my own activities, and performing self-assessment” to “This app provides an efficient way to create visual supports, such as educational, speech-language therapy, and language learning materials.”

The second questionnaire used in this study was adapted from Fontana de Vargas et al., 2022 (see Appendix C). This questionnaire illustrates how the participants perceived the quality of three different areas in the application including: Interaction, Vocabulary Quality, and Overall Usage. Modifications were made to serve the specific purposes of this study’s objectives. First, terminology was adapted across the entire survey from third person (e.g., “Users could easily select a desired vocabulary item within a page”) to first person language (e.g., “I could easily select a desired vocabulary item within a page”). Within the Interaction subsection, two items related to creation of previous communication boards were eliminated as this did not pertain to the objectives of this research (i.e., Users tended to access/use vocabulary from previously created pages, Users tended to access/use vocabulary from newly created pages). Within the “Vocabulary” subsection, one question was modified from “The generated vocabulary included words users did not want to use” to “the vocabulary generated included words I would not have thought of that are relevant.” In addition, three items were added. Specifically, the items added were “The order the vocabulary was presented was adequate,” “The vocabulary generated included words I would target during educational and/or speech therapy sessions,” and “Overall the vocabulary generated is effective in helping me achieve targeted goals for my use.” Lastly, within the “Usage” subsection, one item was modified. Specifically, “Users were more communicative using the application than they usually are using other AAC tools” was modified to “I created topic-specific displays using this application more efficiently than with other AAC tools.” In addition to the two questionnaires, five open-ended questions related to overall experience and vocabulary generation across the two conditions were utilized.

Design and Measures

A descriptive usability study was completed to evaluate the feasibility of using AI to generate relevant vocabulary for TSDs. This prospective design is consistent with a case series (Kooistra et al., 2009) in that the SLPs were exposed to *QuickPic AAC* with the two AI approaches following the reading of the vignette and the outcomes were monitored with observations and via questionnaire. Two dependent variables were measured within this study: (a) specificity of the vocabulary generated across two AI conditions (i.e., Natural Language Generation [NLG] approach based on de Vargas and Moffatt (2021); and the GPT-3.5 approach) and (b) user satisfaction. The specificity of the vocabulary generated was measured in terms of percentages as follows: (a) vocabulary/icons *kept* by the participant for the final TSP relative to vocabulary/icons originally produced by AI; (b) vocabulary kept for the final TSP but with icons *altered* by the participant out of the total # of

vocabulary kept (alteration may involve the participant choosing a different icon to represent the vocabulary identified or moving the existing icon to a different column in the display); and (c) vocabulary/icons *deleted* by the participant from the final TSP relative to vocabulary/icons originally produced by AI (these two measures are inversely related).

Overall user satisfaction with each condition served as the second dependent variable, measured by two questionnaires as described previously.

Procedures

Demographic and AAC Experience Questionnaire

Upon enrollment in this study, participants completed a questionnaire regarding pertinent demographic information and previous experience with AAC. In addition, two brief questions were asked regarding their perspectives towards the benefits of TSDs and the challenges behind the creation of TSDs.

Tutorial

Participants engaged in a two-part tutorial process. Participants initially were provided a printed out *QuickPic AAC* Reference Guide and were asked to read through the Reference Guide independently to familiarize themselves with the functions of *QuickPic AAC*. Subsequently, each participant individually took part in a live tutorial session led by the examiner, during which each feature in the Reference Guide was demonstrated including: creating a new board, editing a board, editing an individual button, changing an individual button's background color, locating a saved board, customizing "My Album," and tips and tricks to create boards.

Experimental Task

Following the tutorial phase, each participant received instructions to generate two TSDs with *QuickPic AAC* utilizing two separate approaches. Participants were aware the purpose of the study was to determine which approach generated more appropriate vocabulary. The two approaches encompassed the NLG method and the GPT-3.5 model. Participants remained blind to both conditions, and the sequence of conditions were randomized amongst participants to mitigate potential order-related effects. The creation of TSDs under both conditions for all participants were screen recorded. This allowed for data analysis to identify the vocabulary selections deemed relevant by participants across both conditions. Participants were provided with explicit instructions for using the app based on the previously described Vignette to create two identical outputs under the two conditions. Additionally, participants were instructed to determine the settings of the app that best suited the child depicted in the vignette, including the number of items populated within each part of speech (e.g., subjects, verbs, prepositions, descriptors, objects), number of columns available of each part of speech, message bar size, and size of the input photo.

Usability Questionnaires

Following their participation in the creation of two mixed displays, participants individually completed a modified version of the MAUQ and a post-questionnaire. These questionnaires were completed independently either directly after the *QuickPic AAC* experience or submitted to the experimenter no later than 24 hours subsequent to their usage of the *QuickPic AAC* application. The post-questionnaires allowed for participants to provide their experiences of the two conditions facilitated by the NLG-AAC approach and the GPT-3.5 approach.

Data Analysis

Data on the perceived benefits and barriers to creating TSDs (AAC Experience Questionnaire) were analyzed descriptively (the small sample sized precluded statistical analysis) by calculating the number of participants who were in support of statements on benefits and barriers, respectively.

Relevant vocabulary was analyzed using simple descriptive summary statistics for each of the approaches (NLG-AAC and GPT-3.5) in terms of specificity. This includes the range, mean, and standard deviations (SD) of the ratios (i.e., percentages) of the vocabulary kept, the vocabulary deleted, and the icons that were modified. As sample size was small, the data were analyzed using the Friedman nonparametric test for several related samples (Daniel, 1990). This test analyzes data for significant differences among the mean ranks for the dependent variables (i.e., vocabulary kept, vocabulary deleted, vocabulary/icons modified). Significant differences were analyzed using the Wilcoxon Signed Rank Test (Rey & Neuhäuser, 2007).

Data on overall usability was also analyzed using simple descriptive summary statistics for both surveys (MAUQ and post questionnaire) for each of the conditions (NLG-AAC and GPT-3.5), including the range, mean, and standard deviations (SD) of the scores in both surveys. Further analysis was conducted within the post-questionnaire. Item analysis was achieved by calculating means across all eight participants per item. Sub-group analysis was also achieved by calculating means across items within the three subgroups. Finally, thematic analysis on the open-ended questions was conducted to reveal overall usability.

Results

Perspectives on Benefits of and Barriers in Creating TSDs

Users' perspectives on the perceived benefits of and barriers to creating TSDs in general (i.e., without *QuickPic* AAC) was revealed through an analysis of the AAC Experience Questionnaire. Participants responded to perceived benefits of TSDs (Figure 2) and barriers in creating TSDs without *QuickPic* AAC (Figure 3). At a group level, 8/8 (100%) participants agreed with the following benefits of TSDs: (a) facilitates expansion of utterance length, (b) helps with addressing communication goals in sessions, (c) helps with modeling of vocabulary, and (d) increases my client's ability to communicate about a specific topic. Additionally, 6/8 (75%) participants agreed that TSDs increased the fluidity of communicating about a specific topic at hand. In terms of barriers to creating TSDs without *QuickPic* AAC, 8/8 (100%) of participants reported the time it takes to create TSDs was a barrier to include them in sessions. There was more discrepancy related to other perceived barriers including: (a) 3/8 (37.5%) participants found it challenging to create visually appealing TSDs and they were unsure of the organization, framework, and guidelines for creating TSDs to include in sessions; (b) 2/8 (25%) participants reported that it was challenging to identify vocabulary and language to target using TSDs, while 1/8 (12.5%) participants reported that they did not have the resources (i.e., apps, software) to create TSDs.

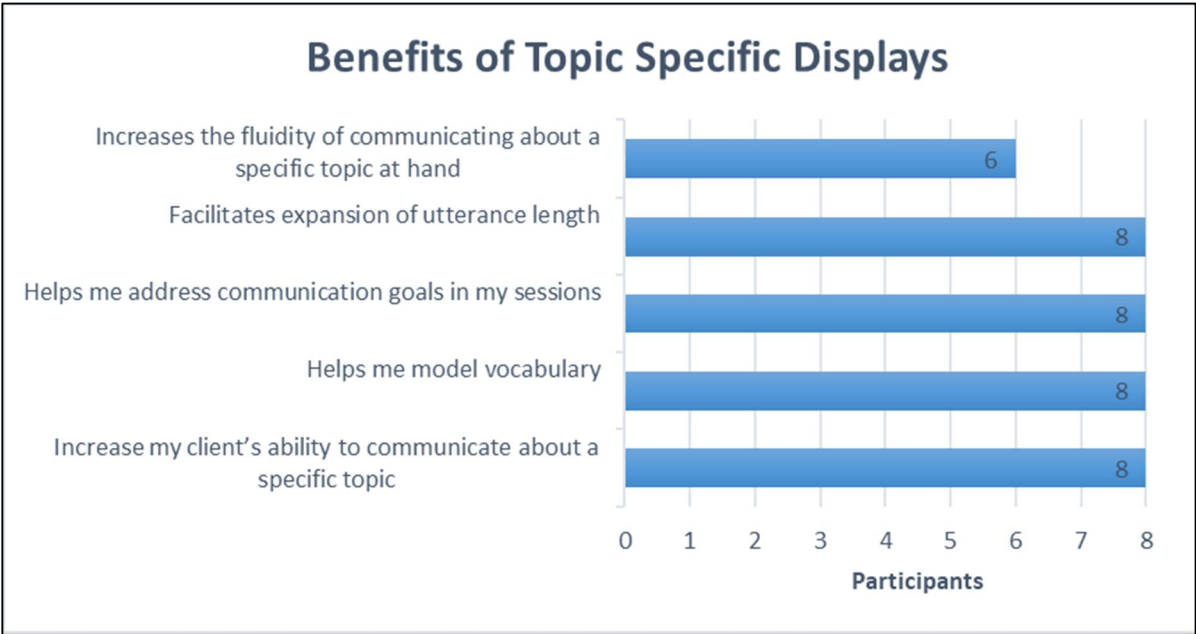


Figure 2. Perceived benefits of topic specific displays.

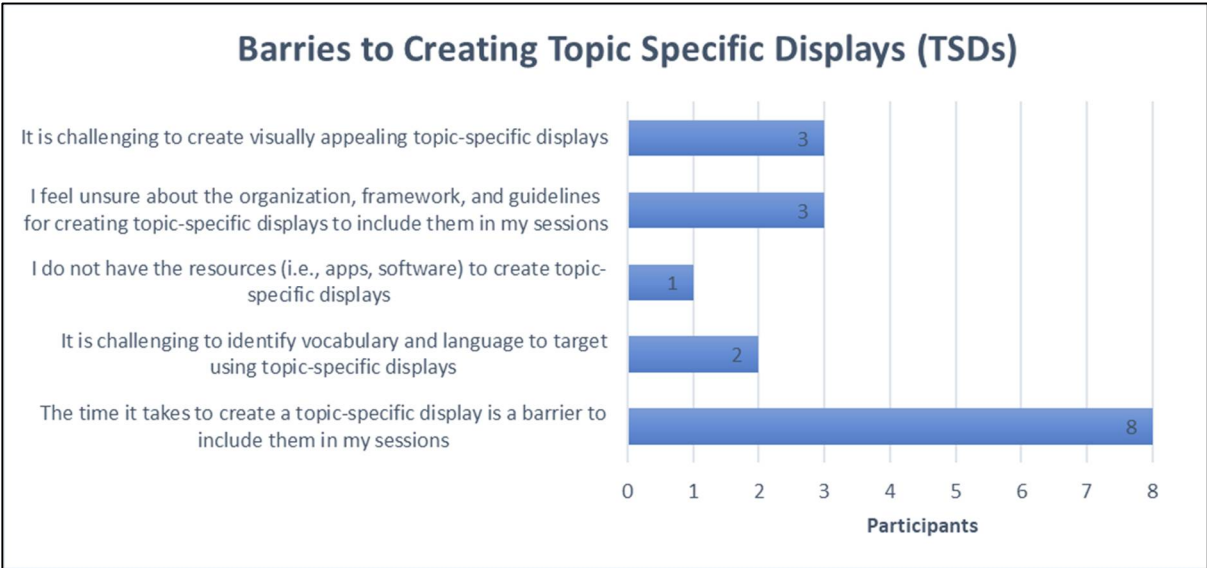


Figure 3. Perceived barriers to creating topic specific displays.

Group-level Descriptive Results

Vocabulary/Icons Kept. Participants were asked to read the vignette and create TSDs under both conditions using the *QuickPic* AAC app. Vocabulary/icons kept by the participants ranged from 6.67% to 64.52% ($M = 38.55\%$; $SD = 20.45\%$) for NLG-AAC and from 33.33% to 100% ($M = 58.04\%$; $SD = 21.89\%$) for GPT-3.5. Across participants (8/8 or 100%), a greater percentage of vocabulary/icons was kept in the GPT-3.5 condition (Figure 4).

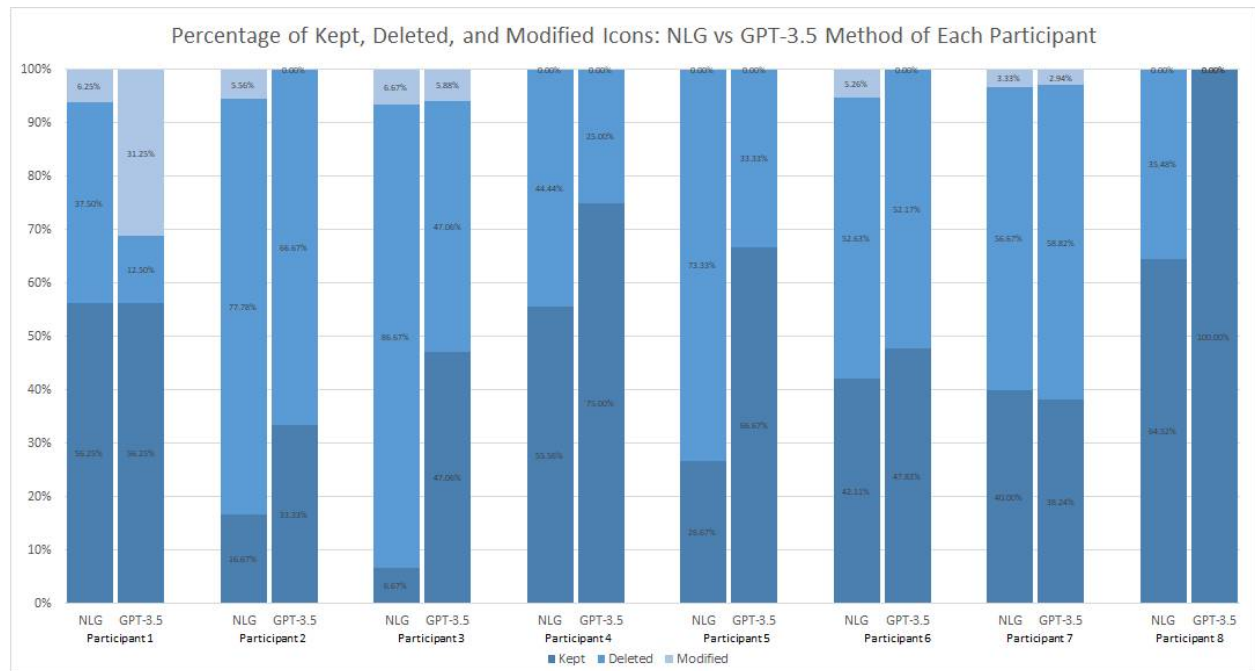


Figure 4. Percentage of vocabulary kept, deleted, and modified by participants across NLG-AAC and GPT-3.5 conditions.

Vocabulary Kept, but with Icons Altered. Some vocabulary was kept by participants but they chose to alter either the icons representing the vocabulary item or they chose to place the icon into a different column of the Fitzgerald Key layout of *QuickPic* AAC. Icons altered by the participants ranged from 0% to 6.25% ($M = 3.38\%$; $SD = 2.97\%$) for NLG-AAC and from 0% to 31.25% ($M = 5.01\%$; $SD = 10.82\%$) for GPT-3.5. Thus, slightly more icons were kept but altered with GPT-3.5.


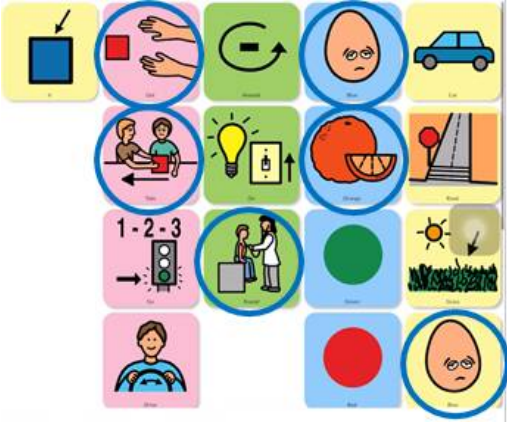
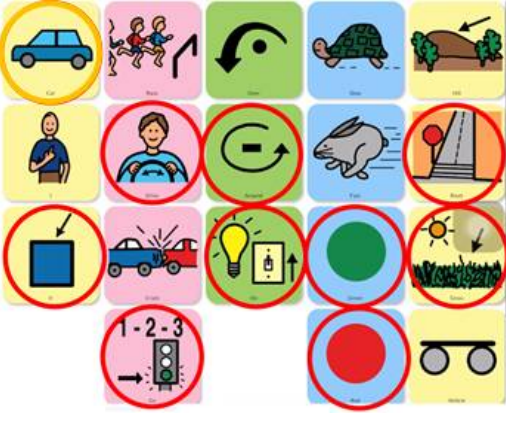
Vocabulary/Icons Deleted. Vocabulary/icons deleted by the participants ranged from 35.48% to 86.67% ($M = 58.06\%$; $SD = 19.23\%$) for NLG-AAC and from 0% to 66.67% ($M = 36.94\%$; $SD = 23.34\%$) for GPT-3.5. Thus, significantly more vocabulary was deleted with NLG-AAC relative to GPT-3.5.

Group-level Inferential Results

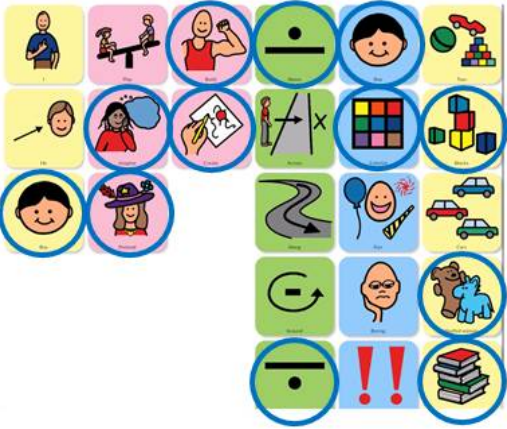

A Friedman Test was conducted to determine if there were statistical differences across conditions (i.e., NLG-AAC, GPT-3.5) among the mean ranks of the vocabulary kept, vocabulary deleted, and the vocabulary modified. A statistically significant differences was found $X^2(5, N = 8) = 26.113$, $p < .001$. This indicates there were differences among the six mean ranks. Three orthogonal contrasts were performed with Wilcoxon tests. For vocabulary kept, the contrasts between NLG-AAC (M rank = 3.88) and GPT-3.5 (M rank = 4.75) were significant ($p < .05$). For vocabulary deleted, the contrasts between NLG-AAC (M rank = 5.13) and GPT-3.5 (M rank = 3.88) were significant ($p < .05$). For vocabulary modified, no significant difference was observed between NLG-AAC (M rank = 1.94) and GPT-3.5 (M rank = 1.44) ($P < .05$).


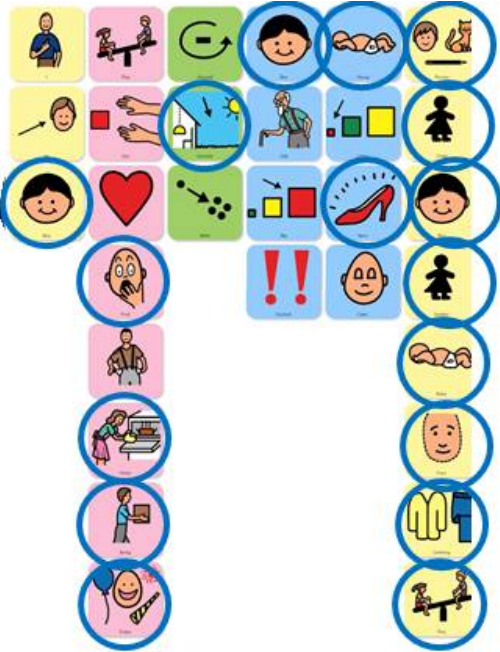

Individual Participant Results

In addition to examining group-level data, it is pertinent to examine participant-level data. Figure 6 displays finalized TSDs per participant created with each condition and the individual vocabulary/icons kept (circled in red), kept but modified (circled in yellow), or deleted (circled in blue).

Participant and Condition	Static Scene Cue Selected	Mixed Displays Generated	
		Original	Finalized
1 NLG			

<div>4</div> <div>GPT 3.5</div>			
<div>5</div> <div>NLG</div>			

<div>6</div> <div>GPT 3.5</div>		
---------------------------------	--	---

<div>7</div> <div>NLG</div>			
-----------------------------	---	--	---

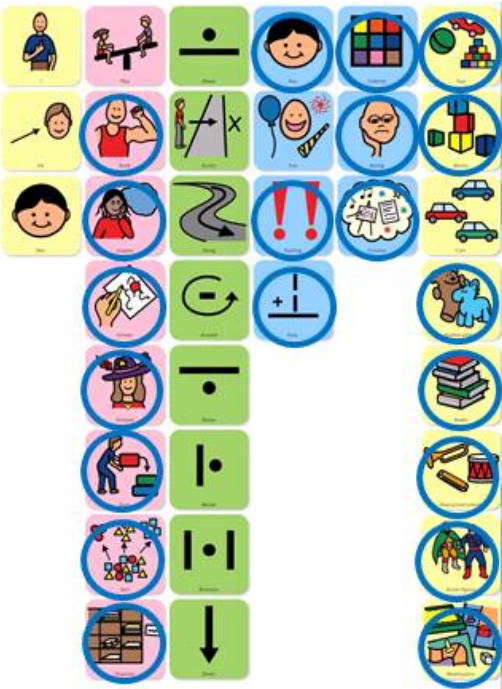
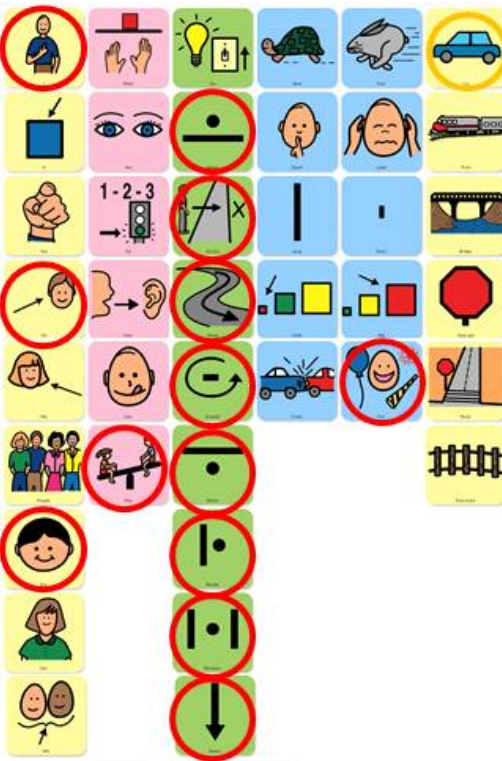



<div>7</div> <div>GPT 3.5</div>			
<div>8</div> <div>NLG</div>			

Figure 5. Comparison of each participant's original and finalized TSDs generated with each condition. Modifications completed by the participant are denoted by the following color-coding: deleted (blue), kept (red), and modified (yellow)

Overall Usability. All eight participants completed two post-questionnaires comparing their experiences between NLG-AAC and GPT-3.5 conditions related to overall experience and satisfaction. Results from the MAUQ are depicted in Figure 6. On a group level, usability scores ranged from 2.43 to 7.00 (M = 4.77; SD = 1.33) and from 4.12 to 7.00 (M = 5.47; SD = 0.86) for the NLG-AAC condition and the GPT 3.5 condition, respectively.

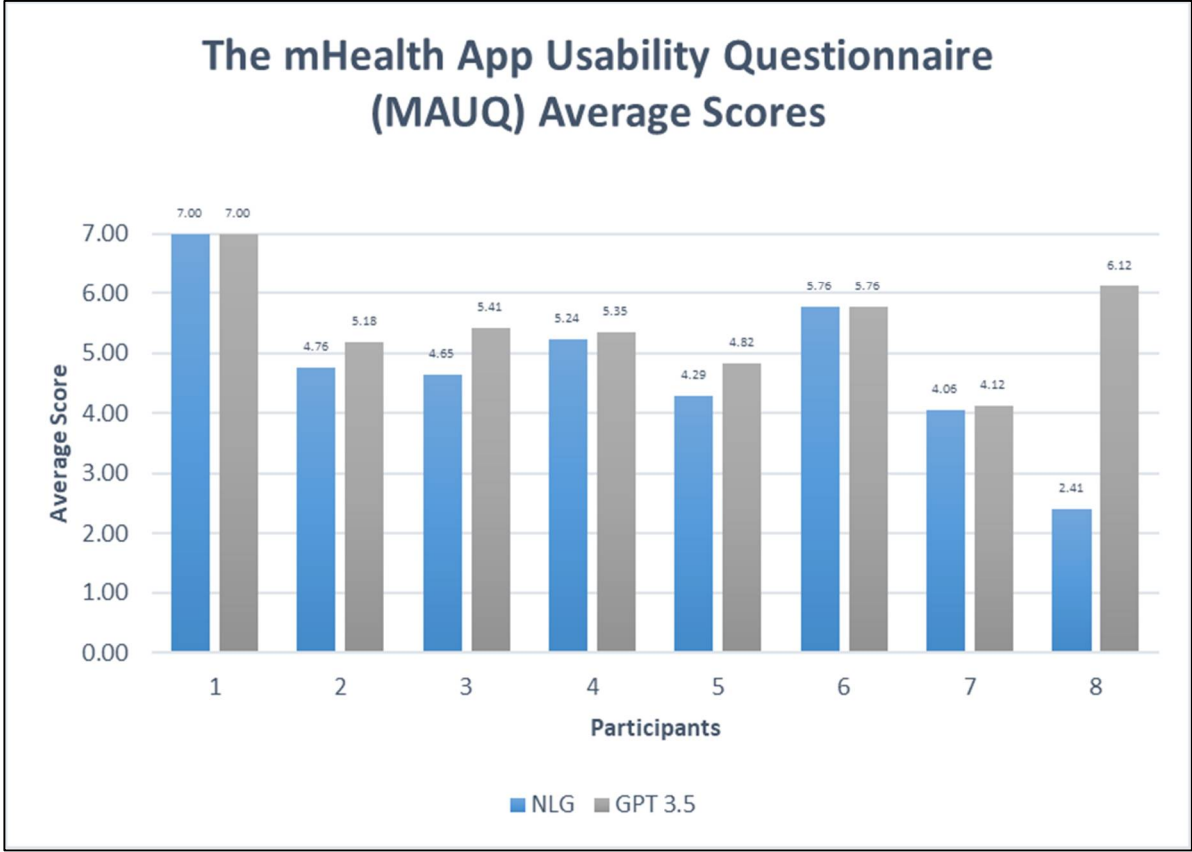


Figure 6. Participant overall average scores comparing the NLG-AAC by de Vargas and Moffatt (2021) to GPT-3.5 using the mHealth App Usability Questionnaire (MAUQ).

The second post-questionnaire participants completed was adapted from Fontana de Vargas et al. (2022). Results from this post-questionnaire are shared in Figure 7. Overall usability scores for the NLG-AAC condition ranged from 3.69 to 5.38 (Mean = 4.80, SD = 0.64) while the GPT-3.5 condition ranged from 4.12 to 6.75 (Mean = 5.82, SD = 0.65). These scores demonstrate overall higher usability scores for the GPT-3.5 approach, reinforcing those obtained complement the MAUQ scores, demonstrating overall higher usability scores on GPT-3.5 condition.

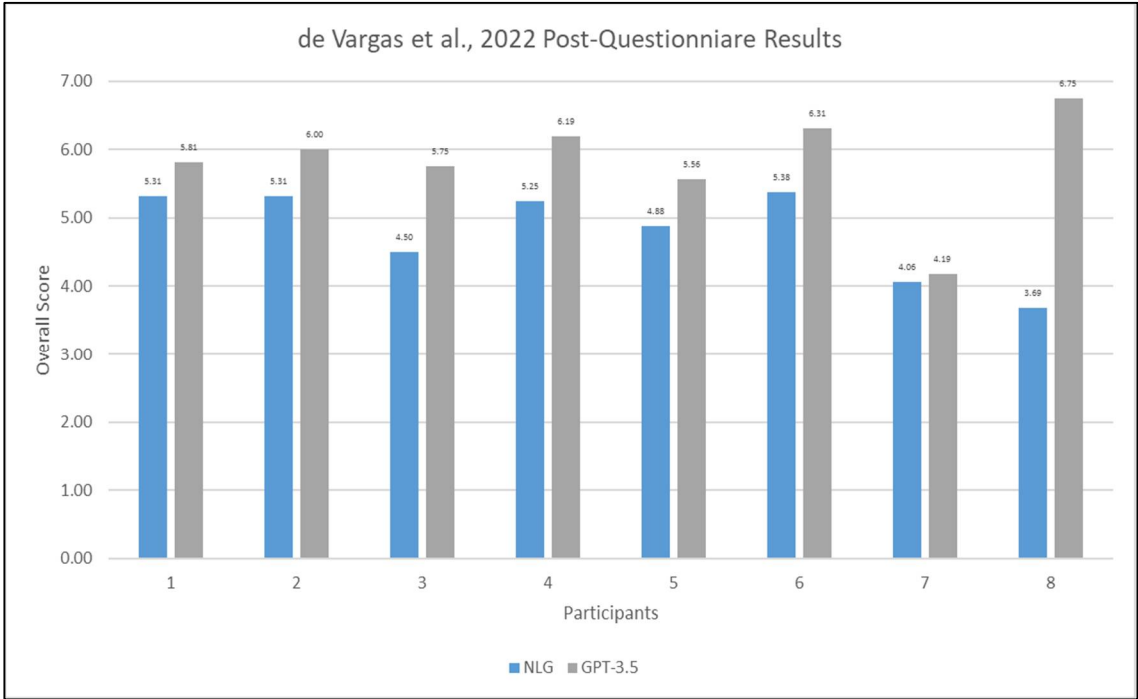


Figure 7. Bar graph demonstrating overall average scores from the de Vargas et al, 2022 Post Questionnaire Survey Results across all participants.

To give a more detailed perspective, the post-questionnaire results were also analyzed at the item level and sub-group level (i.e., Interaction, Vocabulary Generation, and Overall Usage). Figure 8 provides these results in detail. Item analysis was obtained by calculating averages across all eight participants per item. Sub-group analysis was also obtained by calculating averages across items within the three subgroups: Interaction, Vocabulary Generation, and Overall Usage. Most prominently, the Vocabulary Generation sub-group demonstrated the most noticeable difference between NLG-AAC and GPT-3.5 conditions, with an overall greater score in the GPT-3.5 condition.

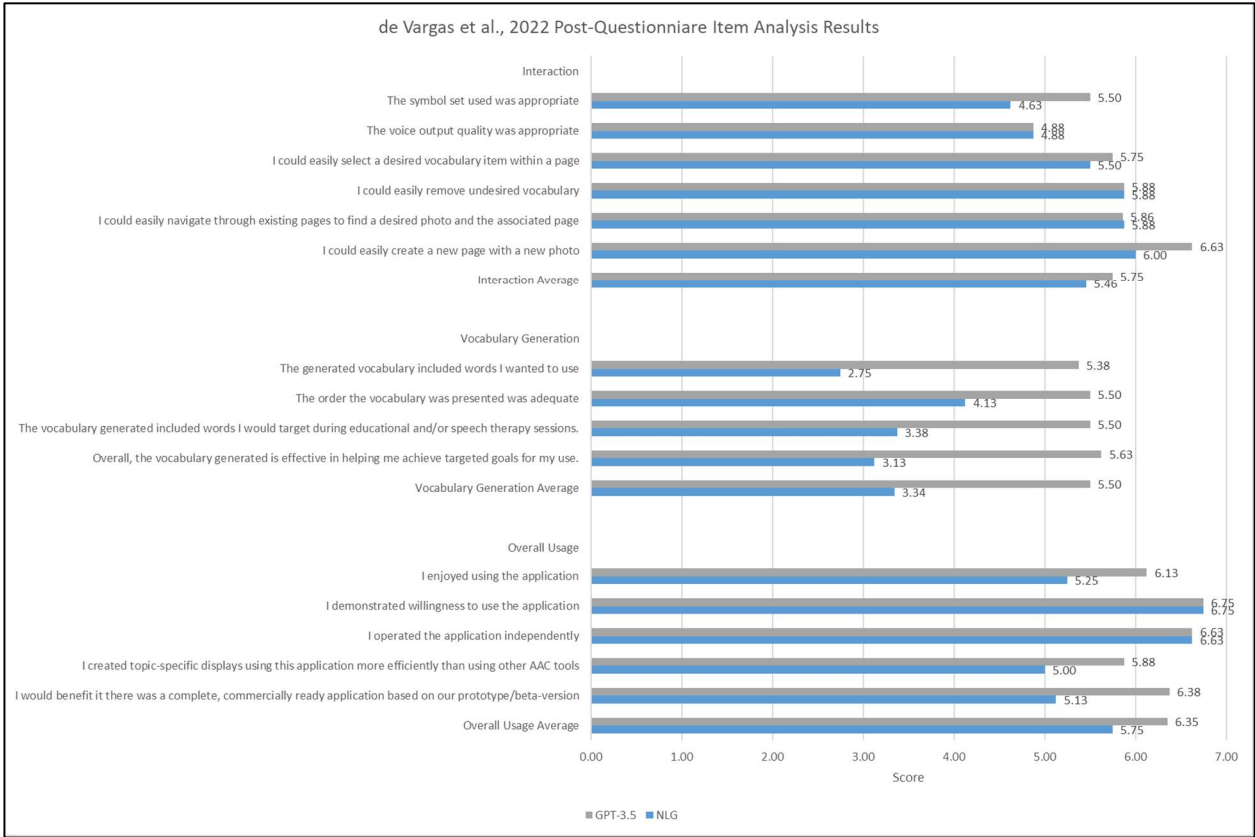


Figure 8. Bar graph demonstrating an item analysis and sub-group analysis from the de Vargas et al, 2022 Post Questionnaire Survey Results across all participants.

Lastly, participants were asked open-ended questions about their experience using *QuickPic* AAC. Results from the open-ended questions on overall experience are presented in Table 1, while use case scenarios are provided in Table 2 from all of the participants. All responses are reported verbatim from the participants, unless indicated otherwise through the inclusion of brackets. Participants were randomly assigned to each condition, and reported on their experience between Experience A and Experience B. The use of brackets was employed to clarify the condition (i.e., NLG and GPT-3.5) being referenced by each participant.

Responses across participants reveal a general unanimous consensus on the feasibility and usability of *QuickPic* AAC in creating TSDs. An overall theme across participant report demonstrated that the app offered a quick and easy way to create TSDs. Notably, two participants even commented that their experience with *QuickPic* AAC surpassed alternative AAC apps (i.e., Boardmaker, TouchChat HD-AAC). Users noted that it was beneficial that *QuickPic* AAC provided a starting point to create TSDs, facilitating the rate at which TSDs could be created. Lastly, users commented on *QuickPic* AAC’s intuitive interface, emphasizing its ease of use and the ease of the editing process. These responses overall demonstrate *QuickPic* AAC’s ability to streamline TSDs.

Table 2. Open-ended questions related to overall experience and vocabulary generation within *QuickPic* AAC across participants.

Describe your overall experience using the *QuickPic* AAC app.

- It was easier to make topic display boards than using Boardmaker or TouchChat HD-AAC. It was faster and helpful that the app provided a starting point.
- Allowed me to easily create a topic specific display.
- Nice quick way to generate topic based displays.
- It was easy to create a topic display based on a simple scene. Editing was simple and effective.
- I enjoyed using the app- the overall learning process felt quick and I felt comfortable navigating and programming it on my own. It was much easier/quicker to program in comparison to another AAC app I have used.
- Great! This updated version is much improved- sleeker with more editing capabilities.
- I like the vocabulary selection feature, but wish I could preview top choices before committing to a specific choice. I felt like the prediction was generic.
- Experience [GPT-3.5] was amazing! So quick and easy to use.

When comparing [NLG-AAC] and [GPT3.5], how do they compare in terms of vocabulary generation and your overall experience?

- I thought [GPT-3.5] generated more appropriate vocabulary and a wider range of appropriate words.
- [GPT-3.5] did a better job of generating vocabulary compared to [NLG]. I needed to change less with [GPT-3.5]
- [GPT-3.5] did a better job. [QuickPic AAC] picked too many irrelevant words which resulted in more time spent deleting
- [GPT-3.5] was significantly better at generating appropriate vocabulary
- [GPT-3.5] generated more appropriate topic-specific vocabulary on its own, so I didn't need to spend as much time editing/programming the page than I did with [NLG]
- [GPT-3.5] included more usable vocabulary- it did include some higher-level vocabulary without some basics (e.g., "imagine" but not "want").
- [GPT-3.5] had more prepositions that I would use. [NLG] had more descriptors I would use, but was missing subject and objects
- [GPT-3.5] was accurate at reflecting words I would want to use. Vocabulary choice for [NLG] was random.

Table 3. Open-ended questions related to use case scenarios and frequency of use across all participants.

Participant	If you had access to QuickPic AAC, would you incorporate it into your practice?	If yes, how?	How often would you use it?

1	Yes	NA	Weekly or monthly depending on my caseload
2	Yes	During therapy sessions to base my therapy on patient's interests	Weekly
3	Yes	With communicators who need explicit support and phrase generation and have trouble navigating across pages	NA
4	Yes	Creating displays "on the fly" in therapy for common activities	Weekly
5	Yes	I would use it to create activity specific topic displays in a much more efficient manner. It would help me increase aided language modeling in sessions.	I regularly see patients who use AAC, so I would use it weekly.
6	Yes	For topic display users and families who are ready to start making their own.	In evaluations on a weekly/daily basis. As a recommendation, every other week.
7	Yes	Help families independently select vocabulary at home	Frequently
8	Absolutely	Creating topic display boards	My patient population is variable. I don't always have patients who need topic display boards. I would use it anytime I needed to create a topic display board.

Discussion

As artificial intelligence (AI) makes significant headway in various arenas, the field of Speech-Language Pathology is at the precipice of experiencing a transformative shift towards automation.

This study aimed to introduce *QuickPic AAC*, an AI-driven application designed to generate topic-specific displays (TSDs) just-in-time from photographs. Specifically, the purpose of this study was to (a) determine which of two AI algorithms (NLG-AAC and GPT-3.5) results in more relevant vocabulary with the *QuickPic AAC* application; and to (b) evaluate the perceived usability of *QuickPic AAC* among practicing speech-language pathologists. The data provides statistically significant evidence that GPT-3.5 consistently generates more relevant vocabulary in that it consistently results in more vocabulary kept for final TSDs and in less vocabulary/icons deleted. It is noteworthy to indicate that the more vocabulary that is kept, the less editing is needed and therefore less time is needed to create personalized TSDs. In general, SLPs expressed an overall high satisfaction in using *QuickPic AAC*. *QuickPic AAC*'s ability to swiftly create user-friendly TSDs may pave the way for other AI-driven tools to enhance language intervention strategies.

A primary focus of this study was on the quality of appropriate vocabulary generated through a specific controlled use case scenario (i.e., vignette). Overall, our findings provided insight that different AI algorithms provide varied vocabulary based on the same stimuli (i.e., photograph). Also, in general, the GPT-3.5 algorithm provided more relevant vocabulary based upon SLPs judgment. A noteworthy discussion point that can be highlighted from the results is the large SD in the percentage of relevant vocabulary kept for both conditions, suggesting a wide variation in the number of vocabulary that participants deemed relevant to keep. Some participants retained a relatively low percentage of icons (i.e., NLG-AAC: 6.67%, GPT-3.5: 33.33%), while others kept a significantly higher percentage (i.e., NLG-AAC: 64.52%, GPT-3.5: 100%). From a clinical standpoint, this poses an interesting finding as it indicates perceived relevance or importance of vocabulary may not be consistent amongst SLPs.

A total of five different photographs served as input stimuli for the eight participants. That means, three photographs (were used twice across two participants each, allowing for a comparison of consistency of each algorithm (NLG-AAC, GPT-3.5). For example, the photo of the "boy playing trains on the train track" was presented to Participants #4 and #8); the photo of "two boys playing race cars" was presented to Participants #3 and #5); and the photo of "the boy in the green shirt playing cars on the wooden floor" was given to Participants # 6 and #7. There was variability in the TSD arrangement (i.e., grid size, number of columns assigned per part of speech, number of icons generated per part of speech) due to participants being instructed to adjust the settings to best suit the child depicted in the vignette. This included differences in grid size, the number of columns assigned per part of speech, and the number of icons generated per part of speech. However, there was no variability in the vocabulary being generated by both algorithms when the same photograph was selected. For example, Participants #4 and #8 both selected the photograph of "the boy in the green shirt playing cars on the wooden floor." In the NLG-AAC condition, Participant #4's settings included up to four icons and one column per part of speech, while Participant #8's settings included up to eight icons and two columns per part of speech. Despite these differences affecting the aesthetics of the TSD, all of the icons generated in Participant #4's TSD were also generated in Participant #8's TSD and in the same order. This was observed consistently across both NLG-AAC and GPT-3.5 conditions, in all three instances. This is further confirmed as Participant #3 and Participant #4 both selected the photo of "two boys playing race cars." In the NLG-AAC condition, both participants' settings were the same (i.e., up to four icons for each part of speech) and the vocabulary generated across both participants were consistent and in the same order.

Because the photographs were the same within each participant (for both conditions) we successfully controlled for threats to internal validity due to item difficulty for within-participant comparisons between the two conditions (NLG-AAC & GPT-3.5)..

Another primary focus of our study was overall satisfaction and usability of *QuickPic AAC* amongst SLP professionals. As discussed previously, the personalized creation of TSDs has a myriad of benefits reported by speech-language pathologists. These advantages include the expansion of utterance length, aiding clinicians in targeting specific communication goals and objectives during sessions, facilitating effective vocabulary modeling, supporting aided language stimulation, and increasing ability to communicate about a specific topic or activity. While the advantages are

apparent, certain barriers were identified in the integration of TSDs, with the largest barrier of all being the time constraints as a primary obstacle in incorporating TSDs into SLP's sessions. Our overall findings demonstrate SLPs were satisfied in using an AI-generated app to create TSDs as it was a quick and efficient way to personalize communication materials for their clients.

Limitations and Future Directions

While our preliminary findings are promising to demonstrate use of AI in speech-language pathology to create TSDs, there are several limitations that need to be recognized. First and foremost, one limitation pertains to the use of different photographs across participants.

With exception of the participant pairs who received the same photos (as described above) they were not kept consistent across all eight participants. Thus, the nature of the photographs may have introduced an extraneous variable that influenced the outcomes. Future research should keep the photographs constant across participants or match the nature of the photos across participants. Relatedly, it is yet unknown whether the nature of the scenes displayed in the photographs afford better or worse AI-powered generation of vocabulary. In the current study, participants used *QuickPic AAC* with only one input photograph across two algorithms. Future research should strive to have participants use multiple input photographs to enhance external validity.

QuickPic AAC allows clinicians not only to delete vocabulary that is deemed "irrelevant" but also to add vocabulary that was not generated by the app. Our procedures did not provide this opportunity for the participants. Hence it is unknown to what degree either of the two algorithms ignored additional relevant vocabulary in the minds of the participants.

Importantly, *QuickPic AAC* is meant to be an additive AAC tool that is used in conjunction with an individual's primary AAC system to enhance communication related to specific topics of interest. Examples of use could be: sharing information about weekend news, describing an activity that occurred at school, discussing a highly preferred area of interest, etc. By bridging the use of *QuickPic AAC* with an individual's primary communication tool, this may provide an avenue for enriched personalized instruction and opportunity to capitalize on teachable moments.

There are some directions in terms of development as well. It is essential to acknowledge a notable restriction of *QuickPic AAC*, specifically its reliance on internet connectivity for functionality as it uses GPT-3.5. This limitation restricts its usage to environments that are equipped with internet access only. This issue highlights the need for future improvement that can enhance its utilization across a broader range of settings. Given our findings that demonstrate GPT-3.5 provides superior relevant vocabulary than NLG-AAC, this study needs to be expanded to include comparison of other AI data sets. Exploring different AI models and their effectiveness in generating relevant vocabulary would likely provide a more comprehensive understanding of the capabilities and limitations of AI being used for vocabulary selection in AAC software and applications. Additionally, at this juncture it would be of value to compare the performance of AI with the performance of humans (i.e., clinicians) in creating topic specific displays. Furthermore, research should examine how *QuickPic AAC* can be implemented in practice settings involving minimally-speaking individuals. Lastly, ethical considerations should be taken into account when integrating AI into AAC practices. Future studies should address issues such as privacy, biases, and security risks.

Conclusions

AI has considerable potential in allied health fields including speech-language pathology. In this study, an AI-driven application designed to generate topic-specific displays from photographs on the fly, *QuickPic AAC*, was evaluated in terms of relevance of vocabulary generated using two different AI-algorithms, and perceived usability. GPT-3.5 produced greater relevant vocabulary compared to NLG-AAC. Additionally, practicing SLPs rated *QuickPic AAC* high in terms of its usability in effortlessly creating topic-specific displays. By embracing AI-technologies such as *QuickPic AAC*, SLPs can leverage its capabilities to alleviate the time demands on creation of personalized materials and dedicate more attention to individualized care and treatment for improving communication skills in individuals.

Author Contributions: The authors made the following contributions: Conceptualization, H.C.S., R.S., C.Y., M.F.V., and L.A.W.; Methodology, R.S.; Software- Programming, M.F.V.; Software- Feedback, H.C.S., C.Y., and L.A.W.; Data collection, C.Y. and L.A.W.; Statistical Analysis, R.K.; Formal Analysis, R.S., and C.Y.; Writing – Original Draft Preparation, C.Y., R.S., R.K. and H.C.S.; Writing – Review & Editing, C.Y., R.S., H.C.S., L.A.W., and M.F.V. All authors read and approved the final manuscript.

Funding: The creation of *QuickPic AAC* and this research received partial support from the App Factory to Support Health and Function of People with Disabilities, funded by a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR) under the U.S. Department of Health and Human Services, specifically, the Shepherd Center (Grant # 90DPHF0004) and Fayetteville Manlius School District.

Informed Consent Statement: Participants provided verbal consent.

Disclosures: This research received partial support from the App Factory to Support Health and Function of People with Disabilities, funded by a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR) under the U.S. Department of Health and Human Services, specifically, the Shepherd Center (Grant # 90DPHF0004) and Fayetteville Manlius School District. The *International Journal of Environmental Research and Public Health* (IJERPH) has granted the authors a no-cost Article Processing Charge (APC) slot for the publication of their manuscript.

Acknowledgements: The authors gratefully acknowledge the participants who generously shared their time and expertise in the development of *QuickPic AAC* and this research.

Conflicts of Interest: Christina Yu, Howard Shane, and Mauricio Fontana Vargas have received a grant from App Factory.

Appendix A. QuickPic Reference Guide provided to participants.

QuickPic
Reference Guide


Quickpic is an app that uses artificial intelligence to create topic specific displays. Topic specific displays are a single page containing all relevant vocabulary to discuss a certain topic/activity. Quickpic allows users to automatically generate relevant vocabulary based off a picture. Users can then select vocabulary to create sentences using synthesized speech.

Create a New Board

1. Select "Create a Quick Pic"
2. Select "Take a Photo," "Upload a Photo," or "Search Web"
 - Take a Photo: uses an integrated camera
 - Upload a Photo: photos stored in your Photos library
 - Search Web: searches for an image via the internet


Edit a Board

1. Select "edit" pencil (bottom left)
 - Swap:** drag icons to another position within the column
 - Trash:** deletes icons
 - Add:** create a new icon
2. **Swap:** Long hold an icon and drag to another position within the column
3. **Trash:** Long hold an icon to delete
4. **Add:** Select + and insert a Label, Message, image, and assign a Part of Speech
5. **Edit:** Long hold an icon and modify the Label, Message, Image, and assign a Part of Speech



Editing an Individual Button

1. Select "edit" pencil (top right)
2. Press and hold the button you want to edit
 - Label:** what is written on the button
 - Message:** what is said using synthesized speech/written in the message window
 - Pronunciation:** how the word is spoken using synthesized speech
3. Select an image using either the "Symbol Library," "Photo Album," or "Search Web"
4. Assign the part of speech for the vocabulary



Changing an Individual Buttons Background Color


1. Select "edit" pencil (top right)
2. Press and hold the button you want to edit
3. Scroll down to "Part of Speech" and assign a different part of speech

Locating a Saved Board

1. Select "My Album" from Home Page

Customizing "My Album"

1. Select "edit" pencil (bottom left)
 - Swap:** drag QuickPics to rearrange My Album
 - Trash:** deletes QuickPics



Tips and Tricks for Use:

1. Choose clear images: Select images that are clear and have objects that are easy to identify. This will make the vocabulary generated more relevant and applicable to the image.
2. Use photographs instead of illustrations: When selecting an image, opt for photographs rather than illustrations. Photographs tend to be easier for QuickPic to understand.
3. Adjust settings to eliminate scrolling: If you want to fit all vocabulary on a page and avoid scrolling within a created QuickPic, make the following adjustments in the Settings:
 - a. Increase the number of columns: Adjust the setting to increase the number of columns available per Part of Speech.
 - b. Decrease the size of the message bar: Adjust the setting to reduce the size of the top message bar. This will free up more space for the image and vocabulary
 - c. Decrease the size of the main photo: Adjust the setting to decrease the size of the main photo. This will allow more content to fit on the screen without requiring scrolling.

Appendix B. mHealth App Usability Questionnaire (MAUQ) for Standalone mHealth Apps Used by Healthcare Providers

mHealth App Usability Questionnaire (MAUQ) for Standalone mHealth Apps Used by Healthcare Providers

[illegible]

.			AGREE
12	Overall, I am satisfied with this app.	<input type="checkbox"/>	DISAGREE <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> AGREE
13	The app would be useful for my healthcare practice.	<input type="checkbox"/>	DISAGREE <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> AGREE
14	The app improved my access to delivering healthcare services.	<input type="checkbox"/>	DISAGREE <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> AGREE
15	The app helped me manage my patients' health effectively.	<input type="checkbox"/>	DISAGREE <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> AGREE
16	This app has all the functions and capabilities I expected it to have.	<input type="checkbox"/>	DISAGREE <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> AGREE
17	I could use the app even when the Internet connection was poor or not available.	<input type="checkbox"/>	DISAGREE <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> AGREE
18	This mHealth app provides an acceptable way to deliver healthcare services, such as accessing educational materials, tracking my own activities, and performing self-assessment.	<input type="checkbox"/>	DISAGREE <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> AGREE

Appendix C. AAC Practitioner/Caregiver's Feedback Questionnaire created by Fontana de Vargas et al., 2022.

AAC Practitioner/Caregiver's Feedback Questionnaire

Based on your experience using our application with your clients/family members, please indicate to what extent you agree or disagree with the following statements:

A. Interaction

1. The symbol set used was appropriate

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
-------------------	----------	---------	-------	----------------

☐

☐

☐

☐

☐

2. The voice output quality was appropriate

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Users could easily select a desired vocabulary item within a page

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. Users could easily remove undesired vocabulary

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Users could easily navigate through existing pages to find a desired photo and the associated page

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Users could easily create a new page with a new photo

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Users tended to access/use vocabulary from previously created pages (e.g., previous days)

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. Users tended to access/use vocabulary from newly created pages (e.g., instants or minutes after creating)

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

B. Vocabulary quality

9. The generated vocabulary included words users wanted to use

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. The generated vocabulary included words users did not want to use

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. The order the vocabulary was presented was adequate

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

C. Usage

12. Users enjoyed using the application

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

13. Users demonstrated willingness to use the application

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14. Users operated the application independently

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

15. Users were more communicative using the application than they usually are using other AAC tools

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

16. Users would benefit it there was a complete, commercially ready application based on our prototype/beta-version

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

References

Daniel, Wayne W. (1990). "Friedman two-way analysis of variance by ranks". *Applied Nonparametric Statistics* (2nd ed.). Boston: PWS-Kent. pp. 262–74. ISBN 978-0-534-91976-4.

Fang, H., Gupta, S., Iandola, F., Srivastava, R.K., Deng, L., Dollár, P., ... & Zweig, G. *From captions to visual concepts and back*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; IEEE: Boston, U.S.A., 2015; pp. 1473-1482. DOI: 10.1109/CVPR.2015.7298754

Frey, B.J. & Dueck, D. (2007). Clustering by passing messages between data points. *Science (AAAS)*, 2007, 315 (5814), 972-976.

Goossens', C., Crain, S. & Elder, P. *Engineering the pre-school environment for interactive, symbolic communication: 18 months to 5 years*. Birmingham, AL: Southeast Augmentative Communication Conference Publications * clinician Series, U.S.A, 1992.

Goossens', C., & Crain, S. Establishing multiple communication displays. In *Augmentative communication: An introduction*, S. Blackstone (Ed.); Rockville, MD: American Speech-Language-Hearing Association, U.S.A., 1986; pp. 337-344.

Greenbaum, T.L. *The handbook for focus group research*, 2nd ed., Thousand Oaks, CA: Sage Publications Inc, U.S.A., 1988

Huang, T.H., Ferraro, F., Mostafazadeh, N., Misra, I., Agrawal, A., Devlin, J., ... & Mitchell, M. *Visual storytelling*. In Proceedings of the North American chapter of the Association for Computational Linguistics: Human Language Technologies; Association for Computational Linguistics, San Diego, U.S.A., 2016, pp. 1233-1239.

Kooistra, B., Dijkman, B., Einhorn, T. A., Bhandari, M. How to design a good case series. *The J of Bone & Joint Surgery* 2009, 91(Supplement_3), pp. 21-26, May 01, 2009. | DOI: 10.2106/JBJS.H.01573

Krueger, R.A., & Casey, M.A. *Focus groups: A practical guide for applied research*, 5th ed; Sage Publishing Inc., U.S.A.; 2014.

Maramba, I., Chatterjee, A., & Newman, C. (2019). Methods of usability testing in the development of eHealth applications: A scoping review. *Inter J of M Inform* 2019, 126, 95–104. <https://doi.org/10.1016/j.ijmedinf.2019.03.018>

- O'Brien, A., O'Brien, M., Schlosser, R.W., Yu, C., Allen, A.A., Flynn, S., Costello, J., Shane, H.C. Repurposing consumer products as a gateway to just-in-time communication. *Sem in Speech & Lang*, 38, 297-312. <https://doi.org/10.1055/s-0037-1604277>
- O'Brien, A., Schlosser, R.W., Shane, H. C., Abramson, J., Allen, A., Yu, C., & Dimery, K. Just-in-time visual supports for children with Autism via the Apple Watch: A pilot feasibility study. *J of Autism & Develop Dis* 2016, 46, 3818-3823. <https://doi.org/10.1007/s10803-016-2891-5>
- Rey, D., Neuhaus, M. (2011). Wilcoxon-Signed-Rank Test. In: Lovric, M. (eds) International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_616
- Schlosser, R. W., Laubscher, E., Sorce, J., Koul, R., Flynn, S., Hotz, L., Abramson, J., Fadie, H., & Shane, H. (2013). Implementing directives that involve prepositions with children with autism: A comparison of spoken cues with two types of augmented input. *Augmentative and Alternative Communication*, 29(2), 132-145. <https://doi.org/10.3109/07434618.2013.784928>
- Schlosser, R. W., Shane, H. C., Allen, A., Abramson, J., Laubscher, E., & Dimery, K. (2016). Just-in-time supports in augmentative and alternative communication. *J of Phys & Develop Dis* 2016, 28, 177-193. <https://doi.org/10.1007/s10882-015-9452-2>
- Schlosser, R.W., O'Brien, A., Yu, C., Abramson, J., Allen, A., Flynn, S., & Shane, H.C. (2017). Repurposing everyday technologies to provide just-in-time visual supports to children with intellectual disability and autism: A pilot feasibility study with the Apple Watch®. *Int J of Develop Dis* 2017, 63, 221-227. <https://doi.org/10.1080/20473869.2017.1305138>
- Shane, H.C., Laubscher, E., Schlosser, R.W., Flynn, S., Sorce, J. F., & Abramson, J. Applying technology to visually support language and communication in individuals with ASD. *J of Autism & Develop Dis* 2012, 42, 1228-1235. <https://doi.org/10.1007/s10803-011-1304-z>
- Shane, H.C., Laubscher, E., Schlosser, R.W., Fadie, H. L., Sorce, J. F., Abramson, J.S., ... & Corley, K. Enhancing communication for individuals with autism: A guide to the visual immersion system. Paul H. Brookes, Baltimore, U.S.A. 2014.
- Soto, G., & Zangari, C. (Eds.). *Practically speaking: Language, literacy, and academic development for students with AAC needs*. Paul H. Brookes, Baltimore, U.S.A. 2009.
- Zapata, B.C., Fernández-Alemán, J. L., Idri, A., & Toval, A. Empirical studies on usability of mHealth apps: a systematic literature review. *J of Med Systems*, 2015, 39, 1. <https://doi.org/10.1007/s10916-014-0182-2>
- Zhou L, Bao J, Setiawan A, Saptono A, Parmanto B, The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. *JMIR mHealth and uHealth*, 2019, 7(4):e11500. DOI: 10.2196/11500. PMID: 30973342

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.