Typically developing preschooler's location of digital photos on speech generating devices using visual scene displays.

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

DANA LYNN THOMPSON

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Joe Reichle, Ph. D.

July 2012

Acknowledgements

Many people contributed to this project and I would like to take the time to thank all of them. I would like to acknowledge Dr. Joe Reichle for helping me come up with the research questions for this project and mentoring me throughout the whole study. I would like to thank Jolene Hyppa-Martin M.A. CCC-SLP for running all of the statistical analyses and giving me feedback when explaining the results. I would also like to thank Laura Nathan, Sheila Cina, and Jackie Hochstein who all did an excellent job collecting the data.

Abstract

Purpose: This study examined preschoolers' (19-25months) accuracy in selecting graphic symbols on the main page of a dynamic visual scene display (VSD) using embedded and non- embedded symbols.

Method: Twenty participants were equally divided into two experimental conditions that included embedded or non-embedded VSDs. Participants were assigned to one of the two experimental conditions and were subsequently taught to select a superordinate symbol on the main page of a symbol display that linked to a second page of each of nine symbols. The participant then selected the target symbol corresponding to the referent displayed at the beginning of the instructional opportunity. After an initial session of nine opportunities participants were provided with up to seven additional intervention sessions. Maintenance was examined subsequent to mastery.

Results: Children in both embedded and non-embedded treatment groups demonstrated improvement over the course of the study. There were no statistically significant differences between children's performance in the embedded and non-embedded VSD conditions. Post hoc analyses revealed that chronological age, receptive language level as measured by the Receptive One Word Picture Vocabulary Test (Brownell, 2000), and gender were significant predictors of participant accuracy. Descriptive analyses revealed some patterns suggesting that participants in the embedded condition made modestly greater gains than participants in the non-embedded condition.

Conclusions: The differential effect of embedded and non-embedded symbol displays was not demonstrated with participants between 19-25 months of age. Based on the current evidence, if VSDs are to be used in an intervention it is likely best for educators to select embedded or non-embedded VSDs based on a comparison of individual learner needs and abilities and to use individual performance data to guide the intervention decision. Limitations of the current investigation are discussed.

Table of Contents

List of Tables

List of Figures

Speech Generating Devices (SGDs) afford children with complex communication needs the opportunity to communicate with others when they have limitations producing speech. To maximize the ease with which beginning communicators can effectively use SGDs, a series of variables must be considered which include: (a) type of symbol display being used; (b) features of the display (e.g. organizational strategy, symbol type, and number of symbols being simultaneously displayed); (c) amount of exposure/instruction to symbols; and (d) child characteristics including chronological age and developmental skills (Olin, Reichle, Johnson, & Monn, 2010). With respect to symbol presentation and display, several investigators have suggested that symbol presentation and the organization of electronic pages may not match the conceptual models that are more consistent with representational and navigational strategies of young children (Drager, Light, Speltz, Fallon, & Jeffries, 2003; Light, 1997). Recently, several investigators have suggested that visual scene displays (VSDs) may be more consistent with young children's organizational skills and are facilitative in initial speech generating device use (Drager et al., 2003; Drager et al., 2004; Light et al., 2004; Olin et al., 2010).The current investigation was designed to compare two variations (i.e., a main page which included embedded target symbols and a main page that did not include embedded target symbols) of VSDs which have been shown to have efficacy with young beginning SGD users (Drager et al., 2003; Drager et al., 2004; Light et al., 2004; Olin et al., 2010).

In this introduction, types of displays and organizational strategies will be described. Next, previous research examining young children's accuracy using different display types in conjunction with SGDs will be discussed. Subsequently, skills acquired by young children that likely facilitate SGD use will be described. Finally, a range of display features (e.g. number of symbols displayed, the type of vocabulary, and treatment dosage) that could affect a young child's performance will be presented.

Display Types Commonly Implemented with Speech Generating Devices (SGDs)

Placing a range of symbols that a child may need to access within the "physical space" available on the pages of a SGD represents a formidable challenge to an interventionist. Typically developing children's vocabulary is acquired at a rapid pace. Paul (2001) reported that by 24 months, children engage in approximately five to seven communicative acts per minute using words and word combinations. The ability to talk about absent objects also emerges around 24 months (although most of children's communication is still grounded within the immediate environment [Paul, 2001]). 24-month-old children have a mean expressive vocabulary size of approximately 312 words (Paul, 2001). Additionally, two-year-old children produce a variety of two word semantic relations (e.g., agent-action, agent-object, action-object), several types of WH-questions (e.g., who? what? where?), and several different grammatical morphemes (Paul, 2001). Consequently, symbol configuration on a SGD for a beginning communicator must carefully address a variety of communicative parameters. Placing a wide range of symbols in an easily accessible display for a young child, increasingly, has been a focus of SGD research in recent years.

Formats commonly used to display graphic symbols on SGDs include static, dynamic, or hybrid (combining elements of static and dynamic) displays. A static display presents all symbols within a learner's communicative repertoire so that all symbols can be seen simultaneously. An advantage of a static display is that it relies on recognition memory because all symbols can be seen at the same time. Unfortunately, static displays limit the number of symbols that are concurrently available to the learner. Dynamic displays offer the advantage of having fewer symbols displayed at a time while still allowing for a larger vocabulary set through linking pages. For example, a food button may be touched on the main page of a display which links the learner to an entirely different page filled with food symbols. Using a scanning selection technique Mizuko, Reichle, Ratcliff, & Esser (1994) demonstrated that having fewer symbols on a page increased young children's accuracy of symbol selection. Although a dynamic display decreases the number of symbols displayed on a single page, it likely increases demands for recall memory in navigating to a specific symbol that is linked to a symbol on the main page of a display. To date, no studies have directly compared static and dynamic display utility with very young typically developing children or children with developmental disabilities who are candidates for graphic mode communicative alternatives.

The Influence of Chronological Age and/or Developmental Status on Performance with Different Symbol Displays and Symbol Organization Schemes

In a traditional dynamic grid-based display, individual symbols are represented inside separate boxes of a grid, organized into rows and columns. In a traditional grid display, symbols are decontextualized without preserving the proportional similarity between the real referent and symbol or even between the symbols themselves (e.g., a line drawing of an apple used to represent an apple may be as big as the head of the boy used to represent boy [Light & Drager, 2007]). An alternative is a dynamic contextual VSD where messages are embedded under “hot spots” (i.e., invisible links to related messages, pages, or symbols) in real world scenes. For example, a digital picture of a park might be displayed. The message, “Push me please” could be selected by touching the swing on the display. Selecting a swing depicted in a park scene is a context with which children are familiar with and may be easier for children than selecting a swing depicted in a box as it would be in a decontextualized grid display.

Some investigators have suggested that young children may perform better with dynamic displays organized into VSDs (Drager et al., 2003; Drager et al., 2004; Light et al., 2004). Drager & colleagues (Drager et al., 2003; Drager et al., 2004; Light et al., 2004) conducted a series of research studies with typically developing children, ages 2 ½-5 years, to investigate the learning demands of different system layouts and language organizations of dynamic display SGDs. Children located symbols within a play context of a birthday party organized in four system approaches: (a) symbols in a grid format organized taxonomically (i.e., symbols organized into hierarchical categories), (b) symbols in a grid format organized schematically (i.e., symbols organized according to events or experiences), (c) symbols in a VSD organized schematically (i.e., a contextual scene with language concepts embedded under “hot spots”), and (d) (for the 4- and 5-year-olds only) iconic encoding (i.e., an encoding technique in which line drawings that are rich in semantic associations are used in combinations as codes to retrieve single words or phrases).Within the grid conditions the main page consisted of four link buttons with single symbols each representing an entire page of related symbols. For example, a picture of a ribbon was used to represent the "opening presents" vocabulary page. In the contextual scene condition, the rooms of the house where the birthday party was being held were used as links from the main page. To access symbols within the "opening presents" visual scene the child selected the living room on the main page where children could be seen opening presents. Drager et al. (2003) found that the 2 ½ -year-old children performed poorly in all conditions but were significantly more accurate in locating symbols in the VSD condition when compared to the grid layouts. Drager et al. (2004) found consistent significant results with three-year-old children. Light et al. (2004) concluded that four and five-year-old children located vocabulary using the VSDs and the grid layouts with similar levels of accuracy.

Results from these studies suggest that, during the early phases of intervention, when compared to grid displays, VSDs may be slightly more efficient for children younger than four years of age. Drager et al. (2003) suggested several possible explanations for their results favoring VSDs with young beginning communicators. Embedding vocabulary within a contextual VSD may reduce metalinguistic demands because symbols were presented in naturally occurring contexts as opposed to isolated decontextualized symbols in a grid display (Drager et. al, 2003). Finding a symbol within a grid display may place a large demand on working memory because a learner must remember the location of many specific symbols (Drager et al., 2003). VSDs may decrease working memory demands placed on young children because they are able to chunk the vocabulary items together that would be used in that thematic scene (Drager et al., 2003). Light and Drager (2007) added that, VSDs organize language which is reasonably congruent with how young children represent vocabulary (Lund, Millar, Herman, Hinds, & Light, 1998). Finally, VSDs preserve the conceptual and visual relationships between symbols (e.g., location, proportionality of concepts) (Light & Drager, 2007).

Visual Scene Options: Embedded and Non-Embedded Displays

In the current investigation, two different strategies to configure a dynamic VSD were examined. *"Embedded"* VSDs involve the selection of a symbol from a main page where target symbols are actually part of the visual scene. When a target symbol is selected a link is activated to a second page that displays a "blown up" image of the portion of the visual scene chosen on the main page. For example, the child sees a picture of the inside of a toy store through glass windows on the main page. When the child selects the toy store, that store will enlarge to fit the screen so the child can see the individual objects in the toy store more clearly.

"Non-embedded" VSDs allow the child to select a symbol from a main page that does not contain the actual representation of the referent that constitutes the actual item that the learner wishes to select. When a child selects this "related symbol" it links to a second page containing the actual symbol corresponding to the referent that is the focus of the communication act which can then be selected by the learner to communicate a message. For example, the child sees a picture of the outside of several stores. After selecting a scene from the main page representing a store, the SGD displays a second page representing the inside of the store where symbols can be selected. To date, studies

have focused on the first display type (Drager et al., 2003; Drager et al., 2004; Light et al., 2004; Olin et al., 2010) and no research has systematically investigated the effects of the different types of VSDs on the navigational skills of young children. In order to use either an *"Embedded"* or *"Non-embedded"* dynamic VSD; matching-to-sample likely represents a critical skill. A matching-to-sample task occurs when an array of choices is offered from which the learner makes a selection that corresponds to a sample (usually a referent in the environment). This referent can be thought of as the instructional cue in the matching task. Typically developing children around two years of age are able to successfully participate in identical matching to sample tasks (i.e., the correct choice is identical to the sample) and some non-identical matching to sample tasks (i.e., the correct choice is not identical to the sample) (Johnston, Reichle, Feeley, & Jones, 2012).

In an embedded VSD, the sample is the actual item that the learner wishes to communicate about and the choices can be viewed on the main page of the VSD. Once a choice is made on the main page, the SGD links to a second page in which a segment of the scene chosen is a "blown up" version of the portion of the main page scene selected. For example, a learner might see a hamburger (i.e., the sample) on a McDonald's® menu and decide that is what s/he wants. S/he then selects the McDonald's® visual scene on his or her communication device where s/he can see a small menu including a hamburger on the main page. Selecting this embedded VSD would result in a transition to page two. An embedded visual scene requires engagement in non-identical matching-to-sample (real desired item to graphic representation). However, although the symbol is two-dimensional because it is embedded, it is identical to the symbol to be located on the second page of the display.

A non-embedded VSD may require higher-level non-identity matching skills because the learner must engage in a non-simultaneous (sample and choices are displayed at the same time) non-identical matching in that the target symbol is not displayed on the main page of the VSD at the point that a sample is made available to the learner (this is referred to as associative matching). For example, suppose that a learner decided that he wanted a hamburger for lunch. The learner would need to select a superordinate symbol of the McDonald's® Golden Arches from the main page. Selecting this symbol would result in a transition to a menu on page two where the learner could select a hamburger symbol to communicate to place his order. A non-embedded VSD is hypothesized to be more difficult for a beginning communicator because the choices on the main page (e.g., the superordinate symbol of the McDonald's® Golden Arches) are very perceptually different from both the sample (e.g., a hamburger) as well as the symbols (e.g., the 2D picture menu) on the second page. Selecting the McDonald's® golden arches which links the learner to a picture menu with a hamburger on it, is presumed to be more difficult than choosing a scene where a picture menu with food items on it can be viewed from the main page. Using a non-embedded VSD is hypothesized to require a child to have more advanced symbolic representational skills; therefore a discussion of how and when children acquire these skills is important.

Representational Challenges for Learner in Utilizing VSDs

Although literature to date suggests the potential efficacy of VSDs, they are not without representational challenges for young children. To effectively use a VSD a child must have acquired the ability to detect and connect symbol-referent relations. Children's symbolic representational skills are not completely developed by the age of two (DeLoache, 2004) which is a challenge in successfully using a dynamic VSD. In a series of investigations, DeLoache & colleagues (DeLoache, 1987, 1995a, 1995b, 2004; DeLoache & Burns, 1994) investigated young children's ability to use a scale model as a representation of a real room in order to locate a Snoopy® doll which provided insight into the developmental process of how children are able to apply representational skills.

In a study by DeLoache (1987) 2 ½ -year-old children viewed a doll being hidden in the scale model and then were told that they would be able to find Snoopy® in the exact same place in a big room but they were unsuccessful in the task. DeLoache (1987) concluded that the 2 ½ -year-olds were unsuccessful because the task required multiple skills. The children had to remember that the little room was a concrete object itself but that it also was a representation of the big room. Investigators reported that the children were interested in the scale model and liked playing with it which made it difficult to see it as anything other than itself. The authors tested this hypothesis by showing the 2 ½ -year-olds a photograph of Snoopy® being hidden instead of using the scale model. As predicted, the 2 ½ -year-olds were much more successful because even though the photograph was highly iconic it was less enticing and salient than the scale model. Thus there may be contexts in which photographs are more instructive as symbols than three-dimensional items serving as symbols. It appears from this work that VSDs are not too representationally complex for young typically developing children.

Other Variables that May Influence Dynamic Visual Scene Display

Other potential variables that could influence the successful use of a VSD are the vocabulary types represented in a display, the number of symbols being simultaneously taught, and treatment dosage (Olin et al., 2010). In a study by Drager et al. (2003) 2 ½ -year-old participants were given an opportunity to learn 12 symbols (six abstract and six concrete) representing vocabulary items and an additional 12 symbols (six abstract and six concrete) representing vocabulary items to learn during the maintenance session. Drager et al. (2003) reported that none of the participants were able to locate abstract items during the initial session and over half did not learn any abstract items during learning sessions. More research is needed to determine the effects of the type and number of vocabulary items has on beginning SGD users proficiency in learning to use a dynamic VSD.

Another important variable that can influence symbol acquisition is the amount of instructional exposure required for individuals (likely as a function of their chronological age and developmental status) to select the appropriate symbol(s) to communicate a message. In Drager et al. (2003) and Drager et al. (2004) children only received four learning sessions and one maintenance session. These authors reported that, overall, participants were not very successful. Olin et al., (2010) examined children's navigational skills using dynamic VSDs to explore child characteristics that may relate to navigational skill differences with even younger children than previous studies. Two groups of participants' ages 24-27 months and 33-36 months participated in up to 12 learning sessions to locate nine symbols representing concrete vocabulary items corresponding to thematic rooms in a dollhouse using a two-page dynamic VSD. These investigators assessed the number of opportunities required to reach an arbitrarily defined performance criterion of accurately selecting the correct symbol during 75% of the consecutively presented opportunities during three consecutive sessions. Results showed that at initial opportunity, older participants were more accurate and significantly faster selecting symbols than their younger counterparts. Younger participants required a greater number of sessions to achieve the performance criterion than the older participants. However, the authors suggested that this difference was likely not clinically significant.

Olin et al. (2010) also matched older and younger participants based on their vocabulary comprehension pretest scores consisting of general vocabulary comprehension scores from the Peabody Picture Vocabulary Test-III (Dunn & Dunn, 1997) or the Receptive One-Word Picture Vocabulary Test (Brownell, 2000). When matched, no performance differences were found during the maintenance session. Thus, comprehension performance may have been associated with the chronological age performance differences reported.

Summary

Available evidence tends to suggest that embedding symbols within a VSD is a more effective organizational strategy than grid displays for toddlers who appear to be readily able to navigate dynamic displays (Drager et al., 2003; Drager et al., 2004; Light et al., 2004). To date, few studies have implemented VSDs with children who are two years of age and younger. Additionally, no research has systematically compared the effects of embedded and non-embedded VSD schemes on the navigational skills of young children. Because the latter strategy has been associated with creating a taxonomic system of storage and retrieval of symbols it becomes relevant to explore the relative difficulty of the two symbol display approaches. Consequently, the current investigation represents an effort to compare young children's performance on these two transition strategies using VSDs in a dynamic page display. Specifically, questions of interest in the current investigation included: (a) Does the performance of children (ages 19-25 months) vary between embedded and non-embedded symbol displays? (b) Does the performance of children (ages 19-25 months) increase with repeated exposures of the VSDs, and if so, (c) Does it differ between the two experimental conditions?

Methods

Participants

Forty-five participants were screened for participation in the current investigation. 17 participants failed a screening and eight refused to participate in the study which resulted in 20 qualifying for participation. The twenty participants were typically developing children between the ages of 19-25 months of age and were recruited from Minneapolis Minnesota area daycare centers. Each child met inclusion criteria that included: chronological age, age appropriate vocabulary comprehension with a standard score of at least 85 on the Receptive One Word Picture Vocabulary Test (ROWPVT) [Brownell, 2000], age appropriate pointing, normal (or corrected to normal) vision, normal hearing, typical cognitive, motor, emotional and language development (as reported by parents and/or teachers), English as their primary language, and consent from legal guardians (a description of screening procedures will follow).

Each of these twenty participants was assigned to one of two experimental conditions by counterbalancing their assignment to obtain a balance of similar chronological ages and standard scores obtained from an administration of the ROWPVT (Brownell, 2000). Statistical analyses were performed to demonstrate that the two groups were relatively balanced for ROWPVT score (p= .01, z= 2.6) and age (p= .02, z= 2.4). Vocabulary comprehension performance yielded a group mean standard score of 96.4 (range=88-110) in the non-embedded condition and 98.1 (range= 86-124) in the embedded condition on the ROWPVT (Brownell, 2000).

The non-embedded condition consisted of five males and five females with a mean age of 21.5 months while the embedded condition had seven males and three females with a mean age of 22.2 months. Participant characteristics are displayed in Table 1.

Screening Procedures

Vision and fine motor screening. Each participant successfully participated during seven out of nine opportunities to match identical line drawings from a modified *Child's Recognition and Near Point Test* (Allen, 1957). The nine line drawn symbol cards were scanned and imported into the DynaVox^{TM} VMax (Series 5) SGD and displayed in a 3x3 grid format. Each line drawing was approximately 1.5 inches x 1.0 inch in size and placed adjacent to one another. Additionally, the nine line drawn symbol cards were scanned, copied, and cut out into flashcards in the same size (1.5 inches x 1.0 inch) as the symbols displayed on the DynaVox™ VMax (Series 5) SGD.

Each participant was given one practice opportunity per line drawn symbol (e.g., bear, horse, truck, cake, telephone, tree, flower, the letter E, and a house) to ensure that the child understood how to perform the vision and fine motor task. The interventionist sat beside the child, held up a symbol, and said, “This is a (symbol name), touch the same.” If the child did not respond within eight seconds or touched the wrong symbol, the interventionist pointed to the correct symbol and repeated the spoken cue, “Touch the same.” If the child still did not respond within eight seconds or touched the wrong symbol, the interventionist guided the child’s hand to the correct symbol while repeating the spoken cue, “Touch the same.” This procedure was repeated for each of the nine line drawn symbols before the child moved on to the screening task.

During the screening, the interventionist stood 15 feet from the participant and randomly held up one flashcard at a time while saying, “Touch the same.” If the participant touched the identical symbol on the DynaVox™ VMax (Series 5) SGD the response was scored as correct. If the child did not respond within eight seconds or the speech output of the DynaVox™ VMax (Series 5) SGD did not match the symbol the response was scored as incorrect. The participant passed the vision screening if s/he correctly touched seven out of the nine symbols displayed. The participant was judged to have sufficient fine motor skills to participate in experimental sessions if s/he could point to the corresponding 1.5 inches x 1.0 inch symbol on the screen because it required appropriate pointing skills and range of motion in order to select a small symbol on a screen with eight other adjacent symbols. The vision screening was modified for one participant who was unable to pass the screening from 15 feet away but passed the screening at five feet. This participant was allowed to continue in the experimental sessions in that the computer screen would be placed directly in front of her and less than five feet away. Eight children failed this vision and fine motor screening task and were not included in the study.

Hearing screening. A Maico Hearing Instruments (Model # MA-25) portable audiometer was calibrated according to the *American National Standard Specifications for Audiometers* (ANSI S3.6-1969 R1989). Participants were taught to raise their hand when they heard a 1000 Hz tone presented at 50 dB HL approximately ten times. If the participants did not reliably respond by raising their hand within five seconds of tone presentation, the interventionist modeled play audiometry (e.g. placing a token in a box contingent on presentation of a tone). When the participants successfully participated in the task for three consecutive opportunities the screening continued with thresholds at 25 dB HL unilaterally at 500, 1000, 2000, and 4000 HZ. Twelve participants would not tolerate wearing the headphones and/or were unable to participate in play audiometry. In these cases, experimenter observations and parental report of normal hearing allowed the participants to advance to experimental sessions.

Vocabulary comprehension screening. Vocabulary comprehension screening was implemented using Receptive One Word Picture Vocabulary Test (ROWPVT) [Brownell, 2000]. To advance to experimental sessions, participants were required to achieve standard scores of 85 or higher. Participants achieved a mean standard score of

96.4 (range=88-110) in the non-embedded condition and 98.1 (range= 86-124) in the embedded condition. Four participants did not meet this criterion and were not included in the study.

Assessment for additional inclusion criteria. Areas addressed via parent and/or preschool/daycare teacher interview included primary language spoken in the home, cognitive and/or emotional delays, fine and/or gross motor delays (not evaluated in the screening task), speech and language delay, and any previous experience with touch screen computers or augmentative communication devices. Based on parental and/or teacher reports each participant spoke English as their primary language. No participants were reported to have any known emotional, cognitive, fine motor, or gross motor delays. Additionally, all parents and/or teachers reported that their child was typically developing in their speech and language development. Additionally, none of the participants were reported to have any experience with touch screen computers or speech generating devices. Two children were excluded from the study because their teachers were concerned about their speech and language development.

Vocabulary pretest of experimental stimuli. To ensure that participants could identify graphic symbols representing vocabulary items used in the current investigation a vocabulary pretest was given to each participant. Each of the nine target vocabulary items (i.e., book, bear, truck, shoes, hat, socks, cookies, juice, and apples) were depicted on separate flashcards. Participants were instructed to point to the picture of each target vocabulary item after the interventionist said its name. The flashcards which depicted target vocabulary items were placed on a table in front of the child using a 3x2 row-column format. The spoken cue “point to/touch the (target word)” provided the child an opportunity to point to the corresponding picture choice. If a participant produced an incorrect selection or did not respond, the interventionist recorded it as an error. Any items incorrectly identified were replaced in a 3x2 row-column format from which the child could identify them using the same procedures and prompts described above. This continued until the child could correctly identify all items s/he previously misidentified prior to proceeding to the experiment. This was done to help ensure that comprehension performance did not influence experimental outcomes. Participants advanced to the action word pretest if s/he correctly identified seven out of the nine target vocabulary items.

After successfully completing the noun pretest, the flashcards depicting a ball, a pair of shoes, and an apple were placed in front of the participant in order to assess comprehension of the action words used in the study. Participants were instructed to listen and point to the target picture. The interventionist gave a spoken prompt for each item, (i.e., “point to/touch something you can eat,” “point to/touch something you can wear,” and “point to/touch something you can play with”). If a participant produced an incorrect selection or did not respond, the interventionist recorded it as an error, pointed to the correct choice while restating the prompt, and then retested after providing the participant with one opportunity to identify each picture corresponding to an action category (i.e. eat, wear, and play). A misidentified action word was replaced in the same array of three (i.e., a ball, a pair of shoes, and an apple) until the child could correctly identify the correct flashcard given the same prompt which included an action word and procedures described above before advancing to experimental sessions. This was done to help ensure that verb comprehension did not influence participants' performance when selecting the correct scene on the main page in the experimental sessions. Participants advanced to experimental sessions after they identified seven out of the nine object names and pointed to the corresponding pictures that matched two out of the three actions (i.e. eat, wear, and play). Three children failed this vocabulary pretest and were not included in the study.

Setting

All sessions took place in the child's home or daycare setting in areas separate from other children. The participant sat beside an interventionist at a table where the DynaVox™ VMax (Series 5) speech generating device (SGD) was placed directly in front of him or her.

Materials

All experimental sessions were structured around dynamic VSDs depicting three businesses including a toy store, GapKids® clothing store, and a McDonald's® restaurant. In the embedded VSD experimental condition the participants could see inside the store where target symbols were displayed on the main page (see Figure 1). In the non-embedded VSD condition participants only saw the outside of the buildings and could not see target symbols on the main page (see Figure 2). Page two was identical in both experimental conditions and displayed a close up of the inside of each business where all target and distractor symbols could be seen. The toy store scene (see Figure 3) included three target symbols (i.e., book, bear, and truck) as well two distractor symbols (i.e., ball and cat) displayed on toy shelves. In the GapKids® clothing store scene (see Figure 4) three target symbols (i.e., shoes, socks, and hat) and two distractor symbols (i.e., pants and shirt) were displayed on shelves and racks. In McDonald's® scene (see Figure 5) three target symbols (i.e., cookies, apples, and juice) and two distractor symbols (i.e., ice cream and hamburger) were displayed on a picture menu board.

The stimuli in each scene were captured by a Samsung™ 12.2 megapixel digital camera. These pictures were downloaded into a computer where Adobe Photoshop ® was used to alter and create the high resolution scenes. The scenes were then imported and displayed on the DynaVox™ VMax (Series 5) SGD. The SGD produced speech output utilizing The AT&T Natural Voices™, Mike 16 (SAPI4), throughout the entire investigation. Additionally, across all sessions, the speech output volume as well as the speech rate remained at 40 (on a 100 point scale). The participants confirmed that the speech volume and rate were both adequate.

The MacArthur-Bates Communication Development Inventories [CDI] (Fenson et al., 1993), a composite list of initial object names that are comprehended by 60% of 19-month-olds, was used to select the nine target vocabulary items. Using the MacArthur-Bates Communication Development Inventories [CDI] (Fenson et al., 1993) to select vocabulary allowed experimenters to ensure that the vocabulary chosen were appropriate for the participants' age and developmental level (see Table 2). The only exception to the 60% criterion were the words play and wear which participants had to comprehend in order to select the correct scene on the main page in both experimental conditions.

However, comprehension of these words was ensured in the vocabulary pretest for each participant. On page two in both experimental conditions (see Figures 3, 4, and 5) all nine vocabulary items represented what could be purchased in a real toy store (i.e., book, bear, and truck), a GapKids® clothing store (i.e., shoes, socks, and hat), or at McDonald's® (i.e., cookies, apples, and juice). In each scene two distractor symbols (i.e., symbols that participants were not asked to locate but were displayed to make the task harder) were also displayed totaling five symbols per scene.

Independent Variable

Embedded and non-embedded main page VSDs constituted the independent variable. In the embedded condition the target vocabulary item was included as part of the symbol displayed on the initial scene displayed on the main page of the display (see Figure 2). During the "non-embedded" condition the target vocabulary item was not included as part of the symbol displayed as the initial scene on the main page of the display (see Figure 1). Additionally, the effect of repeated exposure with feedback and reinforcement served as an independent variable.

Dependent Measures

Accuracy served as one dependent measure in this study. To be considered an accurate response the participant was required to select both the correct scene given the prompt "Touch where Tony can get something to eat." The experimenter changed the final word of the spoken prompt to "wear" or "play with" depending on the target vocabulary item. Once the learner transitioned to page two the speech output had to match the target symbol that the interventionist asked child to find before a predetermined eight second response time ended. The experimenter scored a response as incorrect if the scene selected on the main page did not match the corresponding action category (i.e. eat, wear, or play with), if the speech output on page two did not match the target word, or if the child took longer than eight seconds to respond. The second dependent variable was the number of learning opportunities it took for a participant to reach an arbitrarily defined mastery criterion correctly selecting the main page plus the target symbol on the second page with 75% accuracy during nine consecutive opportunities across three consecutive experimental sessions.

Procedures

Familiarization session. Each child participated in one familiarization session prior to engaging in the experimental sessions. The experimenter presented a scene that was not included in experimental sessions. The scene was set in a kitchen with bananas on the counter. Milk and cheese was placed on top of a table and were clearly visible. Cheese, bananas, and milk were all reported to be words that 70% of 19-month-olds know according to the MacArthur-Bates Communication Development Inventories [CDI] (Fenson et al., 1993). A stuffed animal (Tony the Tiger) sat next to the DynaVox™ VMax (Series 5) SGD and was introduced as the main character using a standard script. "Hello (participant's name), this is Tony the Tiger. Tony the Tiger uses this computer to talk. We are going to help Tony say some words by pressing pictures on the computer." "Tony wants something to drink; can you help Tony say Milk?" If the child responded inaccurately a second prompt was given (e.g. "Touch milk"). If the child touched the corresponding symbol then the response was scored as correct but if the child still did not respond accurately the interventionist guided the child's index finger to the correct picture and scored the response as incorrect. This procedure was administered for all target symbols (i.e., cheese, milk, and bananas). Participants met criterion to move on to the experimental sessions if they correctly touched two out of the three food items. No children were excluded from the study as a result of failing this screening.

Ten participants were assigned to the embedded symbol scene condition while the remaining ten participants were assigned to the non-embedded symbol scene condition Regardless of assigned experimental condition, each child began the experimental sessions with one of five randomized lists of the nine target symbol's names to minimize order effect and to ensure that each list was used equally.

Experimental sessions. Each child participated until reaching the overall mastery criterion (correctly selecting the main page plus the symbol on the second page with 75% accuracy during nine consecutive opportunities across three consecutive experimental sessions). Correct selections occurred when the participant selected the correct scene on the main page and the symbol on page two. Sessions were held two to eight days apart (M=4).

In the non-embedded VSD the participants viewed only the storefronts of the businesses (Toy Shop, GapKids® clothing store, or McDonald's®) on the main page (see Figure 1). On the main page of the embedded VSD the specific target symbols could be seen through glass windows (see Figure 2). When a business was touched, the device moved to a linked second page. This second page consisted of a close up of the business selected (Toy Shop, GapKids® clothing store, or McDonald's®) and was identical in both experimental conditions (see Figures 3, 4, and 5).

All opportunities were presented randomly approximately 30 seconds apart. Participants were given one opportunity to locate all nine target symbols per session. Each participant was given a spoken prompt "Touch where Tony can get something to eat." Once the child reached page two of the experimental opportunity a second prompt was given "Touch (target word)." The selection was scored as correct if the scene selected from the main page transitioned to the second page where the target vocabulary word could be selected and resulted in the voice output matching the target word.

If the child made an incorrect symbol choice on the main page (i.e. touching the wrong storefront) an error correction procedure was implemented. For example, if the child was asked to touch where Tony can get something to eat but incorrectly touched the toy store instead of McDonald's® the researcher said, "Nice try! That is where Tony can get something to play with; touch where Tony can get something to eat" while clearly pointing to McDonald's®. If the child did not respond or took longer than eight seconds to respond, the opportunity was scored as incorrect and the experimenter implemented correction procedure just described. If the learner still did not respond within eight seconds the experimenter guided the child's index finger towards the correct scene. When the learner selected an incorrect symbol from the main page s/he was automatically prompted on the second page by the experimenter with verbal prompt, "Touch (target word)" while clearly pointing to the correct symbol on page two. If the child made an error in selecting the target symbol located on page two, the same error correction procedure was used.

Maintenance session. Each participant completed one maintenance session. Participation in experimental sessions ended when the participant completed a total of eight sessions or when the participant achieved the mastery criterion (correctly selecting the main page plus the symbol on the second page with 75% accuracy during nine consecutive opportunities across three consecutive experimental sessions), whichever came first. The maintenance session took place approximately two weeks after the participant's final experimental session. The procedures used in maintenance sessions were identical to those used in experimental sessions.

Procedural Fidelity and Response Reliability

A procedural checklist was created to evaluate procedural fidelity among three interventionists in administering screenings, familiarization sessions, experimental sessions, and maintenance sessions. Interventionists and reliability observers were trained prior to the investigation to a reliability and fidelity criterion of 90% (agreements divided by agreements plus disagreements multiplied by 100) for two consecutive sessions. Independent observers used the procedural checklist to score procedural fidelity on 15% of all sessions by observing a minimum of one session in each condition for 30% of the participants. All procedural fidelity measures were computed as the number of procedural steps completed correctly divided by the total number of procedural steps specified on the trained procedural checklist, multiplied by 100. Procedural fidelity taken during screening and familiarization sessions was 100%. During experimental sessions fidelity was 95% (range= 84%-100%). Fidelity during maintenance sessions was 94% (range=87%-100%). This indicates that all interventionists consistently followed procedural guidelines.

Response reliability was also evaluated on 15% of all sessions by an independent, on-site observer. Response reliability was computed for at least one screening, familiarization, experimental, or maintenance session for 30% of the participants. Response reliability measures were computed as agreements divided by agreements plus disagreements, multiplied by 100. Response Reliability during each screening, familiarization, experimental, and maintenance sessions was 100%. This indicates that the interventionists and independent observers consistently agreed on correct vs. incorrect responses.

Data Reduction and Analysis

Analyses were performed after data were organized on a spreadsheet by participant, scene, session, and VSD page (main page which was either embedded or non-embedded, and second page which was the same in both conditions). Accuracy was computed as the number of correct responses divided by the total number of opportunities, multiplied by 100. Statistical analyses were completed using R statistical analysis software version 2.13.1 (R Development Core Team, 2011).

Accuracy

Main page mean accuracy versus mean overall accuracy. An accurate response on the main page included the participant selecting the correct VSD from an array of three VSDs that were either embedded or non-embedded depending on the experimental condition to which a learner was assigned. A correct overall response included a correct response selecting the correct symbol located on the main page and a correct selection of the specific target symbol located on the second page. Mean accuracy was computed across all opportunities for participants in each experimental condition. Additionally, group mean for response accuracy on the main page versus group mean accuracy in overall performance was computed. These data were compared and descriptively summarized.

To determine if there was a statistically significant difference in the frequencies of correct responses between the embedded and non-embedded performance during maintenance, a Chi-Square statistic with a continuity correction (Yates, 1934) was computed.

To further examine performance accuracy, a mixed linear regression model was applied to examine the number of correct main page responses in maintenance sessions between groups, controlling for initial accuracy in the first session, ROWPVT score and gender or age. Both models treated subjects as random effect (Faraway, 2006).

Mean overall accuracy in the first session, maintenance, and best performance sessions. Mean overall accuracy was computed for participants in the embedded and non-embedded groups. These means were compared for the first session, maintenance, and best session performance. Best session performance was the peak performance for each participant during a session following the first session. These comparisons were made to examine if one group acquired or maintained more navigational abilities than the other.

Mean overall accuracy for each VSD. Mean overall accuracy was also computed and compared for embedded and non-embedded experimental groups for each of the three individual VSDs in the first session and maintenance sessions. This was computed to detect any performance differences that might relate to aspects of the individual VSDs and their corresponding target vocabulary.

Mastery

Number of participants achieving overall mastery. The number of participants achieving overall mastery (correctly selecting the main page plus the second page with 75% accuracy during nine consecutive opportunities across three consecutive experimental sessions) was also computed for both groups. To determine if there was a significant difference in the frequencies of participants achieving overall mastery between the groups, a Chi-Square statistic with a simulated *p*-value based on 2000 replications of the data was computed (Patefield, 1981). Additionally, the total number of sessions required for participants to reach mastery was computed and described for each group.

Mastery of main page versus overall mastery. Main page mastery was calculated by analyzing each participant's performance in correctly selecting the main page with 75% accuracy during nine consecutive opportunities across three consecutive experimental sessions. Overall mastery was achieved if a participant correctly selected the main page plus the second page with 75% accuracy during nine consecutive opportunities across three consecutive experimental sessions. These data were compared to explore differences between groups in terms in mastery of performance on main page only versus mastery of performance across the overall trial.

Results

Performance during Experimental Conditions: Accuracy

Main page accuracy versus overall accuracy at maintenance in embedded and non-embedded conditions. During the first session, group mean accuracy on the main page was compared to maintenance for both embedded and non-embedded groups. Results show that both groups exhibited a similar relationship between main page accuracy (responses on the main page only) and overall accuracy (responses on the main page plus the second page). Figure 6 displays the total percentage of main page accuracy and overall accuracy by group. In both the embedded and non-embedded groups, main page accuracy was better than overall accuracy.

Although not statistically significant, participants in the non-embedded condition were slightly more accurate when calculating overall group accuracy than those in the embedded condition. Overall non-embedded group accuracy (main page plus second page) was 48% (range= 27%-91%), compared to 46% (range= 25%-91%) in the embedded condition. Accuracy for only the main page was 59% (range=36%-95%) in the non-embedded condition, compared to 55% (range=36%-95%) in the embedded condition.

[Insert figure 6]

The Chi-Square statistic was applied to determine if there was a significant difference between accuracy (frequencies of correct responses) in the two experimental conditions during maintenance. Accuracy at maintenance was not significantly different between the two groups (X^2=0.22, df=1, *p*=0.64). To determine if one group's navigational skill acquisition was significantly different from the other a mixed linear regression model was fitted to examine differences between the total number of correct main page responses in maintenance sessions between groups (controlling for initial accuracy in the first session). There was not a significant difference between the embedded and non-embedded accuracy. The test for condition found no effect (*p*=0.74, z = 0.33). However, performance on the ROWPVT and gender were both significant predictors of accuracy in maintenance sessions (*p*= .01, z= 2.6 and *p*= .05, z= 2.0, respectively). Regardless of condition, participants who had higher ROWPVT scores and were male performed better in maintenance (see Figure 7). A second mixed linear regression model controlling for age found no effect for condition, but did find that age significantly predicted accuracy in the maintenance session (*p*= .02, z= 2.4), with older participants achieving higher accuracy at maintenance (see Figure 8).

[Insert Figures 7 and 8]

Group mean overall accuracy collapsed across embedded and non-embedded in the first session, maintenance, and best sessions. Figure 9 displays group mean overall accuracy (main page plus second page) in the first session, maintenance, and best session performance for both the embedded and non-embedded groups. During the first session, group mean performance was less accurate than group mean performance in the maintenance sessions. Thus, both groups acquired navigational abilities during the experiment.

For both of the experimental groups, group mean overall accuracy in the best session was better than group mean overall accuracy in either the first session or maintenance sessions. Neither group retained peak performance during the maintenance session. In the non-embedded condition, group mean overall accuracy in the first session was 46% (range= 11%-89%), best session was 68% (range=44%-100%), and during maintenance was 60% (range= 22%-100%). Participants in the embedded condition had group mean overall first session accuracy of 34% (range=0%-89%), best session accuracy of 72% (range=44%-100%), and maintenance accuracy of 59% (range=22%-100%).

Descriptively, greater gains were evident during the embedded condition when compared to overall accuracy in the non-embedded condition. That is, the embedded participants had both lower mean accuracy in the first session (34%) than the non-embedded participants, and higher mean accuracy in best session (72%) than the non-embedded participants, for a total change in performance of 38 percentage points. Whereas, the non-embedded group had first session mean accuracy of 46%, and best session of 68%, for a total change in performance of 22 percentage points. However, these results did not reach statistical significance (*p*=.21, f=1.67, df=1, 20).

[Insert Figure 9]

Overall group mean accuracy in the first session and maintenance sessions separated for embedded and non-embedded VSD conditions. The first session and maintenance mean accuracy was compared in both the embedded and non-embedded conditions separately for each of the three VSDs to detect any performance differences that might relate to aspects of the individual VSDs and their corresponding target vocabulary. These results are displayed in Figure 10. Overall (main page plus second page) maintenance accuracy was nearly identical for each of the three VSDs in both embedded and non-embedded conditions. Overall group mean accuracy during the first session for each of the main page VSDs ranged from 30% to 60% (M=39%, SD=11.4) and in maintenance sessions group mean accuracy across all VSDs in the first session ranged from 57% to 63% (M=60%, SD=2.3).

[Insert Figure 10]

Mastery

Number of sessions and participants achieving overall mastery. A total of six participants achieved overall mastery (correctly selecting the main page plus the target symbol on the second page with 75% accuracy during nine consecutive opportunities across three consecutive experimental sessions.), four from the embedded group and two from the non-embedded group. For both the embedded and non-embedded groups, 20%

of participants (2 out of 10) achieved mastery on the third session. In the embedded condition, an additional two participants achieved mastery on the eighth session while no other participants achieved mastery in the non-embedded condition. Examination of the frequency of participants achieving overall mastery between the groups via a Chi-Square statistic computed via simulation and the difference between groups for achievement of overall mastery was not statistically significant (X^2=0.95, p=0.61).

Mastery of main page versus overall mastery. When comparing overall mastery (correctly selecting the main page plus the target symbol on the second page with 75% accuracy during nine consecutive opportunities across three consecutive experimental sessions) the embedded group had 40% reach mastery compared to 20% in the non-embedded group. However, an equal number of participants achieved mastery of main pages in each group. Forty percent of participants reached main page mastery (at least 75% accuracy on the main page during nine consecutive opportunities for three consecutive sessions) in both the embedded and non-embedded groups.

[Insert Figure 11]

Summary of Results.

The only statistically significant differences found in participant performance on the independent variable was for the influence of age (*p*= .02, z= 2.4), ROWPVT performance (*p*= .01, z= 2.6) and gender (*p*= .05, z= 2.0). Each of these factors served as predictors of accuracy in maintenance sessions (see figures 6 and 7). Although this effect involved both experimental conditions, gender, age or ROWPVT score did not yield a statistically significant differentiation of performance between the two experimental conditions.

Main page accuracy (accurately selecting the main page only) was compared to overall accuracy (accurately selecting main page plus second page) in both experimental conditions during the maintenance session (see figure 8). Participants in the non-embedded condition were slightly more accurate when calculating overall group accuracy and main page accuracy than those in the embedded condition.

Group mean overall accuracy (main page plus second page) was compared in the embedded and non-embedded experimental conditions during the first session, maintenance, and best sessions (see figure 9). Results showed that both experimental groups acquired some navigational abilities. That is, during the first session group mean performance was less accurate than group mean performance in the maintenance sessions. For both of the experimental groups, group mean overall accuracy in the best session was better than group mean overall accuracy in either the first session or maintenance sessions. Although results did not reach statistical significance, greater gains were evident during the embedded condition when compared to overall accuracy in the non-embedded condition.

Overall group mean accuracy was compared in both experimental conditions during the first session and maintenance sessions for each VSD (see figure 10). Variability between each of the three individual scenes (i.e., McDonald’s®, GapKids® clothing store, and the toy store) in the embedded and non-embedded experimental conditions decreased in the maintenance session in which participants' performance was nearly identical, ranging from 57% to 63% (M=60%, SD=2.3). Overall group mean accuracy during the first session for each of the main page VSDs ranged from 30% to 60% (M=39%, SD=11.4).

Forty percent of embedded participants achieved overall mastery, whereas 20% of non-embedded participants achieved overall mastery. When accounting for only main page mastery the groups had an equal number of participants reaching the mastery criterion.

Discussion

Toddler performance on embedded and non-embedded dynamic VSDs were compared to determine: (a) Does the performance of children (ages 19-25 months) vary between embedded and non-embedded symbol displays? (b) Does the performance of children (ages 19-25 months) increase with repeated exposures of the VSDs, and if so, (c) Does this increase differ between the two experimental conditions?

Previous investigators had demonstrated that organizing graphic symbols into VSDs results in a performance advantage for young children when compared to other organizational strategies, such as traditional grid displays (Drager et al., 2003; Drager et al., 2004; Light et al., 2004). However, VSDs can be organized in different ways and previous investigators focused on using embedded VSDs (Drager et al., 2003; Drager et al., 2004; Light et al., 2004, Olin et al., 2010). A symbol embedded in a visual scene is not necessarily exactly identical to its referent (since the referent could be seen in a variety of contexts), but there is greater perceptual similarity between the referent to be located and the symbol than in a non-embedded VSD. In this latter display, the learner must engage in a non-simultaneous *and* non-identical matching task. Non-simultaneous and non-identity matching are both viewed as more challenging for young learners than simultaneous matching and identity matching (Johnston et al., 2012). Consequently, the current experiment tested the hypothesis that an embedded graphic symbol display would be associated with superior accuracy when compared to child performance learning graphic symbols in a non-embedded display.

Descriptively, several patterns of performance provided somewhat modest support for the hypothesis that embedded VSDs may yield greater accuracy than non-embedded displays. For example, more participants in the embedded group achieved overall mastery than in the non-embedded group. However, the descriptive performance differences reported consistently failed to reach a level of statistical significance.

Answers to Research Questions

The effects of non-embedded vs. embedded main page VSDs on accuracy.

Statistically significant differences were not found between the embedded and non-embedded experimental conditions. The only variables that demonstrated a statistically significant association with accuracy included chronological age, vocabulary comprehension as measured by the ROWPVT (Brownell, 2000), and gender. The two groups were relatively balanced for ROWPVT score and age, but the embedded group had seven males and three females, whereas the non-embedded group had five of each gender. Previous investigations had not reported gender have an influence on VSD performance, but in the present study males tended to perform than females. Eight of the 20 participants achieved accuracy levels in the maintenance session that were 70% or better, and six of those eight participants were male.

As might be expected, participants with higher ROWPVT scores tended to perform better at maintenance. All five of the participants who were 100% accurate in the maintenance session were in the top half of the distribution of ROWPVT scores. Four of those five also had the highest four ROWPVT scores among all participants. With regard to chronological age, the five participants who scored 100% in the maintenance session were also among the oldest in the sample, ranging from 23-25 months of age. The only other investigation to examine the effect of vocabulary comprehension status and age was Olin et al., (2010) and the present results support those findings.

Perhaps with a greater number of symbols displayed on the main page the magnitude of difference in performance between the two treatment groups would have been enhanced. That is, although a ceiling effect was not observed, it is possible that a greater number of symbols displayed on the main page may have increased cognitive load to a level that began to facilitate differentiation between the two conditions.

Additionally, the small number of participants included in this study may have decreased the likelihood of demonstrating statistical significance. With only 10 participants assigned to each experimental condition (particularly given that both experimental conditions were not examined within participant) made it somewhat less likely to obtain statistically significant differences. A comparison of group mean performance in the first session versus best session yielded superior performance for the embedded group with performance improving 16 percentage points more than the non-embedded group. Although not a statistically significant difference, this may warrant further exploration using larger participant groups with more challenging main page displays.

Further descriptive data also suggest that participants in the embedded group did make modestly greater accuracy gains than participants in the non-embedded group. When comparing overall (main page plus second page) group accuracy in the first session, participants performed better in two out of the three non-embedded VSDs. However, a comparison of overall group accuracy during the maintenance session between experimental conditions, participants performed slightly better in two out of the three embedded VSDs. The possible emerging differences at maintenance may have implications for examining maintenance across a longer period of time to see if performance on non-embedded scenes is apt to degrade more readily than performance with embedded VSD displays.

More participants in the embedded group achieved overall mastery than in the non-embedded group. Mastery data from the main page showed that performance of the two groups were equivalent in that 40% of participants in each group reached mastery. However, when computing overall mastery (main page plus second page) mastery remained at 40% in the embedded group but decreased to 20% for the non-embedded group. It is possible that participants allocated so much effort to the main page of the non-embedded scenes that after transition to the linked second page their attention was more apt to wane resulting in poorer performance. A measure of response latency may have served as a measure of more subtle effect of cognitive effort being expended on the non-embedded main page in making a correct selection when compared to symbols displayed in the embedded condition.

The effect of repeated exposures on accuracy between the experimental conditions. Results suggest that children's performance in each of the two experimental groups improved with repeated opportunities. Participant accuracy improved with additional opportunities in both experimental groups. Peak performance for all participants occurred in one of their final three sessions. Given decreases in performance for both experimental groups during maintenance, it is reasonable to conclude that the arbitrary criterion may not have been sufficient. Future investigators may want to consider implementing a more stringent criterion.

During initial sessions participants may not have been attending to the embedded symbols but with increased opportunities learners began to attend and use them. Overall group mean accuracy during the first session was better in the non-embedded condition for two of the three VSDs. However, this pattern reversed for overall accuracy in the maintenance session. In maintenance, participants performed better in the embedded VSD condition for two of the three VSDs. These results suggest that over time participants began to attend and use the embedded symbols. Incorporating eye-tracking technology into VSD studies would allow investigators to determine if participants were attending to the embedded symbols more in later sessions compared to earlier sessions.

Additional Explanations for the Study's Outcomes Based on Previous Research

Iconicity and VSDs. Embedded VSDs could be considered to have greater iconicity to the adults creating them, given the fact that icons representing actual target vocabulary are actually depicted in the VSD. However, these VSDs may not necessarily be more iconic to the 19 to 25 month old participants using them, given the U-shaped learning trajectory reported by Namy and colleagues (2004).

The iconicity hypothesis suggests that symbols that resemble their referent are easier to recognize and learn to use than more abstract symbols (Fuller & Lloyd, 1991; Lloyd & Fuller, 1990). This hypothesis is supported by studies, including children and adults with and without disabilities, which have suggested that the use of iconic symbols might be a better communication choice for people with severe intellectual disabilities or young children because recognition of the relationship between a symbol and referent is easier if there is some degree of perceptual similarity or iconicity (Mizuko, 1987; Mirenda & Locke, 1989; Mizuko & Reichle, 1989).

However, Namy, Campbell, and Tomasello (2004) proposed that iconicity does not have a consistent effect on symbol use for young children. These investigators found that there is a U-shaped developmental trajectory when mapping iconic or arbitrary symbols to their referents. Their study compared children's ability to learn arbitrary versus iconic gestures as symbols during a "finding game" at 18 months, 26 months, and four years. Results showed that 18-month-olds could map both iconic and arbitrary symbols to a referent equally well, 26-month-olds succeeded only in the iconic condition, and four-year-olds were successful with both. They argued that for young children first learning the use of symbols, iconicity is not relevant, and it is only when young children are more competent symbol users that iconicity is helpful for them. The researchers concluded that 18-month-old children may not see the similarity between a referent and an iconic symbol, and thus iconicity does not necessarily provide an advantage for this population or for the current study's participants.

Visual complexity and VSDs. In the current investigation, the embedded VSDs could be considered more visually complex (i.e., depicting multiple objects, shapes, text and/or colors) than their non-embedded counterparts. For example, 100% of the McDonald's® embedded VSD was visually complex, compared to approximately 75% of the non-embedded McDonald's® VSD. The entire embedded GapKids® VSD had complex visual stimuli, compared to only 50% of the surface area of the VSD used in the non-embedded condition. Finally, slightly less than 90% of the embedded Toy Store VSD consisted of complex visual stimuli compared to about 50% in the non-embedded Toy Store VSD. Consequently, it is possible that the increased context provided by the embedded VSDs might have been counterproductive, being distracting rather than supportive.

Previous investigators have found that increasing the complexity of the background of symbols may actually decrease accuracy and reaction times in locating them (Dixon, 1981; Wilkinson & Snell, 2011; Thistle &Wilkinson, 2009). Dixon (1981) showed that the match-to-sample tasks were easier when individual objects were cut out of photos compared to the when the picture of the same object was depicted against a background. Thistle and Wilkinson (2009) found that the addition of background color to a symbol was more distracting than facilitative to typically developing preschoolers. Wilkinson and Snell (2011) demonstrated that when background color that was meant to be facilitative was added to a symbol to signal the valence of the emotion, accuracy in symbol selection actually decreased. These results indicate that in some instances multiple features incorporated into a graphic display that inherently make the display more visually complex may make vocabulary location tasks more difficult.

Therefore, increased visual complexity within the embedded VSDs when compared to non-embedded VSDs might have reduced the anticipated facilitative effect. More research will be needed to determine when the amount of context provided within a VSD becomes counterproductive rather than facilitative.

Variations of the current study's procedures from previous studies. The current investigation differed from previous studies examining young children's use of VSDs in a couple of ways. First, the present study was different from previous reports in that it had an array of only three VSDs for the participant to choose from on the main page. Mizuko et al., (1994) demonstrated that having fewer symbols on a page increased young children's accuracy of symbol selection. It is possible that the small number of visual scene symbols on the main page in the current investigation may have made it a potentially easier discrimination task than in many previous studies investigating VSDs. This, in turn, could have minimized the need for the participant to use the "embeddedness" of the VSD as an instructional cue. Likewise, the current study used a set of nine nouns as target vocabulary while previous studies used a larger number of target vocabulary. However, given that few participants reached mastery, the task used in this investigation did not appear to be particularly easy for participants.

A second variation that could have influenced results was that action vocabulary items were used to cue the participants prior to a response opportunity. One might argue that following an instructional cue containing an action word (i.e., *eat*, *play* or *wear*) may actually be more difficult than navigating to specific nouns as in other studies' procedures. A majority of existing literature has focused on locating nouns, which are more easily represented by icons than are action words. The more iconic representations of the target nouns that were provided in the embedded VSDs used in this study may have had less of a facilitation effect because the corresponding cue did not mention the target noun.

Limitations of the Current Investigation

There were several limitations in the current study that should be addressed which include: a small sample size of typically developing participants, fewer scenes and target symbols to choose from than previous studies, an unequal amount of visual complexity between conditions, and the lack of eye-tracking technology as a dependent measure.

Replications with larger numbers of participants are likely needed to detect statistically significant differences and to improve the external validity of the findings presented here. Given the small sample of children in this study, variables that are more difficult to control such as attention to task or motivation may have contributed to individual performance differences that could have influenced results. The reinforcement value of small tokens offered during the investigation may have lost some of their value as the study progressed thereby masking improvements. Larger samples would lessen the possibility that these variables would influence results.

Children who require AAC often have complex cognitive, physical, and communication needs and may have performed differently than the typically developing children participating in this study. For children with communication disorders to have full benefit of these investigations, this line of research needs to be extended to populations of participants with disabilities.

In the current study, the embedded VSDs were inherently more visually complex given that numerous target symbols were displayed on the main page. The additional context provided in the embedded VSDs could have been distracting rather than facilitative which could have effected participants' accuracy and reaction times when locating the embedded symbols.

The present study differed from previous VSD studies in two ways which may have decreased the facilitative effect of the embedded VSDs. First, a majority of existing literature has focused on locating nouns, which are more easily represented by symbols than action words. The more iconic representations of the target nouns in the embedded VSDs may have had less of a facilitation effect because the corresponding cue did not mention the target noun. Second, the present study had fewer scenes and target symbols than previous studies making the task easier which may have resulted in enhanced learner performance and minimized the need for participants to use the "embeddedness" of the VSD as an instructional cue.

Therefore, it is possible that some participants did not attend to the embedded symbols. Future investigations should consider using eye-tracking technology as a dependent measure to clarify where participants' attention is being directed.

Directions/Implications for Future Research

Even though the literature has documented that VSDs are easier for very young children to navigate when compared to other types of organizational strategies used on dynamic displays, peak group mean performances in the 60% accuracy range for each VSD in this study highlights the substantial room for improvement in VSD performance (and also perhaps in instructional strategies) for very young children. Additional research could explore the influences of the number of main page VSDs, the effects of nouns or action words used in the prompts, and the number of target vocabulary set on the accuracy of vocabulary location by young children using dynamic displays. It could also examine the effects of chronological age and gender (Stephenson, 2009). These explorations may shed light on the cognitive skills required to master a dynamic display VSD and the instructional strategies to teach to most effectively teach its use.

Future studies which utilize eye-tracking technology may be particularly useful in examining where children look when viewing VSDs and how much context and background should be provided to maximize navigation accuracy while not creating counterproductive visual complexity. The Embedded VSDs used in this study were more visually complex than non-embedded VSDs which could have made focusing on and finding the embedded symbols more difficult. Therefore, eye-tracking technology would allow investigators to examine whether participants are attending to the critical element in an embedded scene or not.

Utilizing not only eye-tracking technology but also response latency as an additional dependent measure of potential task difficulty would be important to incorporate into VSD studies. Learners experiencing a greater cognitive load in the non-embedded condition may take relatively longer to respond in that condition. This greater difficulty may have resulted in greater diligence allocated to the main page at the cost of decreased diligence to the linked second page. Response latency represents a measure that may assist in identifying more subtle differences in main page performance challenges for young children.

It may also be beneficial to examine the influence of VSDs on participants with Autism Spectrum Disorder (ASD) and compare their performance with that of typically developing controls matched for chronological and intellectual level. It is possible that children with ASD may be less apt to derive a benefit from embedded symbols if they engage in selective attention. Selective attention has been demonstrated to be more prevalent in this population than typically developing children and at least some other populations with ASD.

Alternatively, participants with ASD may perform better than typical developing controls when using an embedded VSD. The central coherence theory considers an individual’s ability to focus on details as well as whole or gestalt of a situation to extract meaning from context in order to “see the big picture”. Individuals with typical information processing systems show a preference for global information over individual parts (Shah & Frith, 1983). Individuals with Autism Spectrum Disorder are known to have a weak central coherence meaning that they have a preference for parts and may fail to see the whole picture. For example, many children with ASD identify faces by individual parts better than controls (Campbell, Baron-Cohen, & Walker, 1995) but controls do better at interpreting whole facial expressions than individuals with ASD (Deruelle, Rondon, Gepner, & Tardif, 2004). Findings from a study by Ring, Baren-Cohen, and Wheelwright (1999) showed that children with Autism Spectrum Disorder were better able to find embedded objects and text than controls. Thus, future research exploring the addition of context by embedding symbols into a VSD may have different results for typically developing individuals versus those with disabilities such as Autism Spectrum Disorder. This makes an additional case for incorporating eye tracking into VSD studies, as well as including participants with Autism Spectrum Disorder or other developmental disabilities.

Educational Implications

For most young children, locating symbols in a dynamic display is challenging, even with a small number of highly iconic symbols displayed in a contextual VSD. The results of the present study do not provide strong evidence for the superiority of embedded or non-embedded VSDs. Consequently, it is likely best for interventionists to select the VSD type that is most appropriate for a specific learner by performing a within participant comparison of the two strategies.

Children need an AAC system that will grow with them and allow them to increase their expressive language skills. Ideally, the system will allow the child to create messages with all the syntactic, semantic, morphological, phonological, and pragmatic aspects of language. In general, creating novel utterances with these features may be more difficult using only VSDs. To date little attention has addressed this issue. Thus, we do not yet know whether visual scene displays are advantageous for learners who are beginning to acquire these aspects of language.

Summary

The purpose of this study was to assess the effects of two different strategies to configure a dynamic VSD (i.e., embedding and not embedding target symbols on the main page of a dynamic VSD on participants' (ages 19-25 months) ability to navigate and accurately select the correct scene on the main page given a prompt with an action category (i.e., eat, wear, or play) and the target symbol on page two within a two-page dynamic VSD). Results did not yield any statistically significant differences between the groups assigned to each condition, but descriptive analyses revealed some patterns suggesting that learners made more gains in the embedded condition as opposed to the non-embedded condition. Suggestions have been made regarding future research needs involving VSDs. Although this study did have limitations, data presented here is encouraging in that it documents that very young children can achieve meaningful levels of accuracy in vocabulary location tasks on dynamic display devices with VSDs. Based on the lack of statistically significant evidence available at the present time, if VSDs are to be used it is best for educators to select embedded or non-embedded VSDs based on a comparison of individual learner needs and abilities and to use individual performance data to guide the intervention decision.

Table 1.

Participants' chronological age, gender, vocabulary comprehension performance expressed as a standard score, and assigned experimental condition.

Participant #	Age in Months	Gender	ROWPVT	Experimental Condition
1	25	F	107	Non-Embedded
2	20	F	91	Non-Embedded
3	19	M	88	Non-Embedded
4	19	F	99	Non-Embedded
5	23	M	97	Non-Embedded
6	20	M	91	Non-Embedded
7	24	F	97	Non-Embedded
8	23	F	110	Non-Embedded
9	20	M	90	Non-Embedded
10	22	M	94	Non-Embedded
11	24	M	96	Embedded
12	24	F	91	Embedded
13	22	M	94	Embedded
14	24	M	124	Embedded
15	19	F	93	Embedded
16	23	M	96	Embedded
17	21	M	88	Embedded
18	22	M	86	Embedded
19	24	M	120	Embedded
20	19	F	93	Embedded

Note. Receptive One Word Picture Vocabulary Test (ROWPVT) [Brownell, 2000]

Table 2.

Percentage of 19-Month-Olds Comprehending all Symbols Represented in each Visual Scene Reported in the MacArthur-Bates Communication Development Inventories (CDI)

Visual Scenes Displayed			
Symbols Displayed	Toy Store	GapKids® Clothing	McDonald's®
Symbol #1	Book 91.8%	Shoes 89%	Cookies 82.2%
Symbol #2	Bear 60.3%	Socks 71.2%	Apples 74%
Symbol #3	Truck 61.6%	Hat 69.9%	Juice 78.1%
Distractor Symbol #1	Cat 67.1%	Shirt 45.2%	Ice Cream 46.6%
Distracter Symbol #2	Ball 94.5%	Pants 43.8	French Fries 42.5%

Figure 1. Main page of the embedded VSD.

Figure 2. Main page of non-embedded visual scene display (VSD).

Figure 3. Toy store page two in both experimental conditions.

Figure 4. GapKids® clothing store page two in both experimental conditions.

Figure 5. McDonald's® page two in both experimental conditions.

Figure 6. Comparison of overall accuracy (main page plus second page) versus main page accuracy between the experimental conditions.

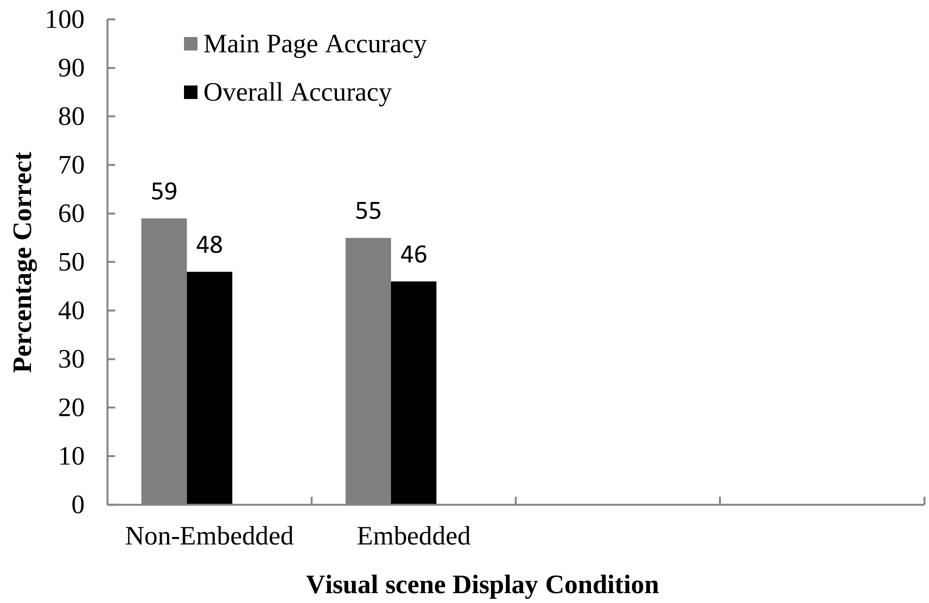


Figure 7. Comparison of accuracy in the maintenance session between the experimental conditions when participants are ordered by their Receptive One Word Picture Vocabulary Test standard scores.

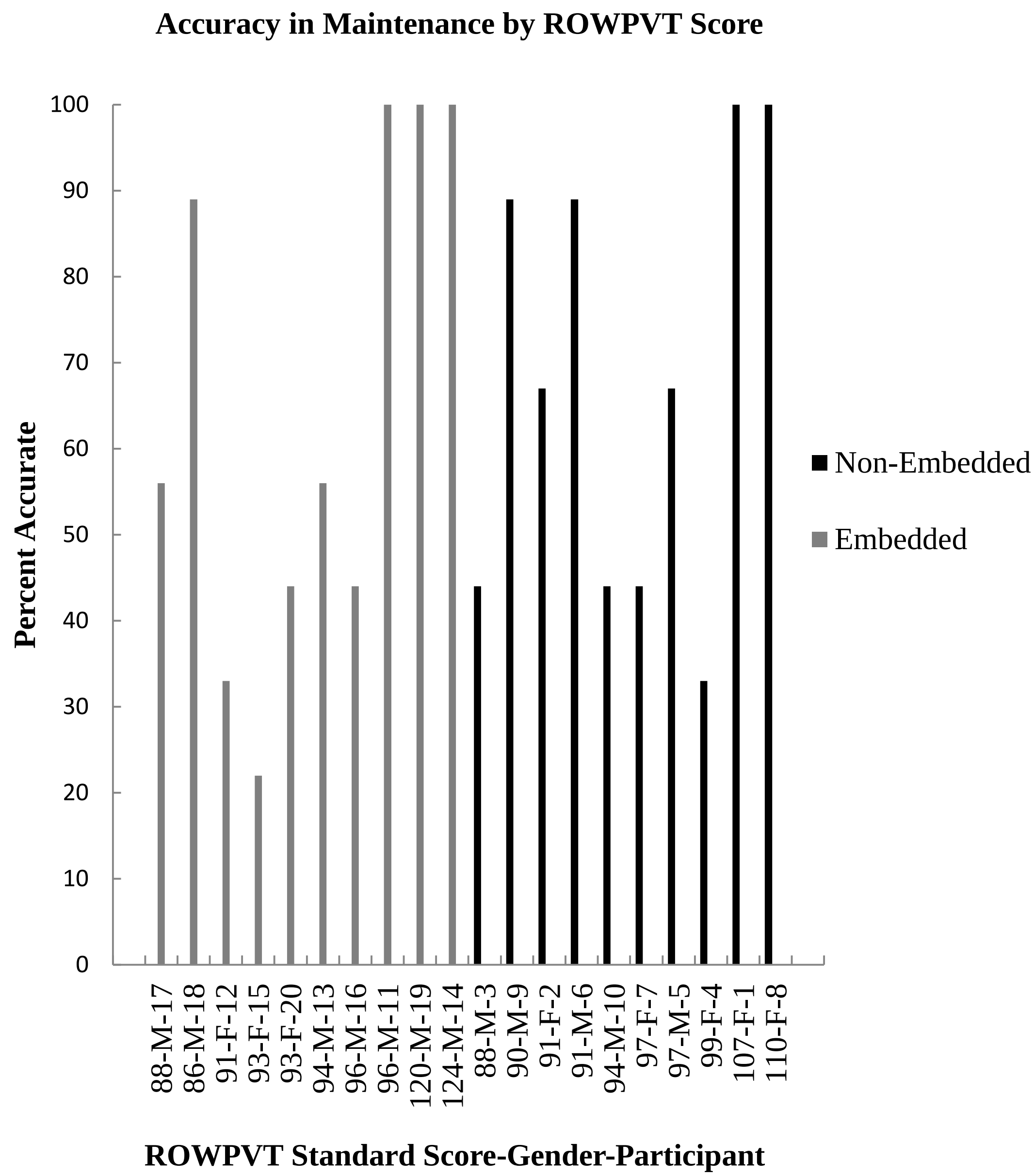


Figure 8. Comparison of accuracy in the maintenance session between the experimental conditions when participants are ordered by age from youngest to oldest.

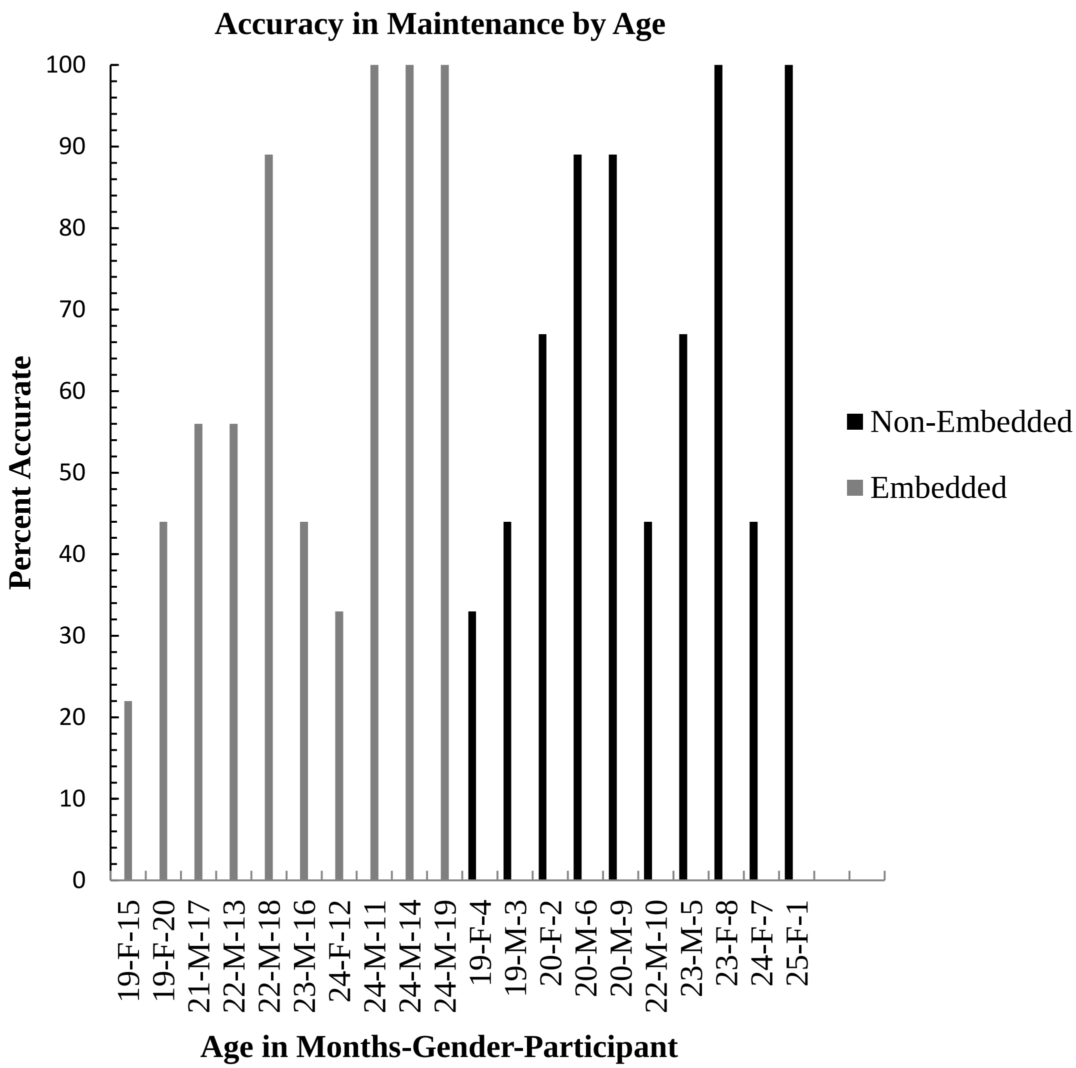


Figure 9. Participants overall accuracy (main page plus second page) in embedded and non-embedded conditions during the first session, best session, and maintenance.

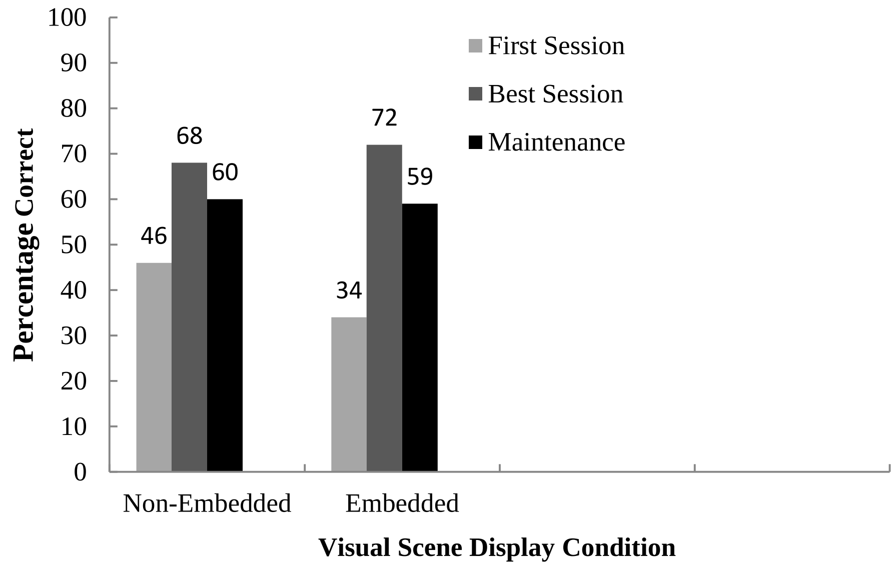


Figure 10. Overall accuracy (main page plus second page) for each VSD during the first session and maintenance sessions in each experimental condition.

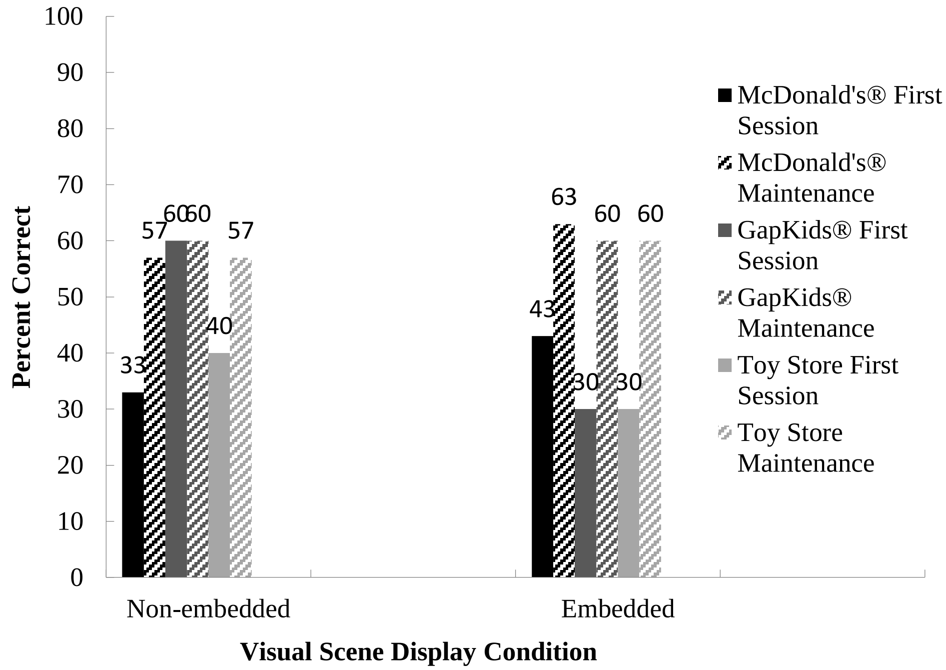
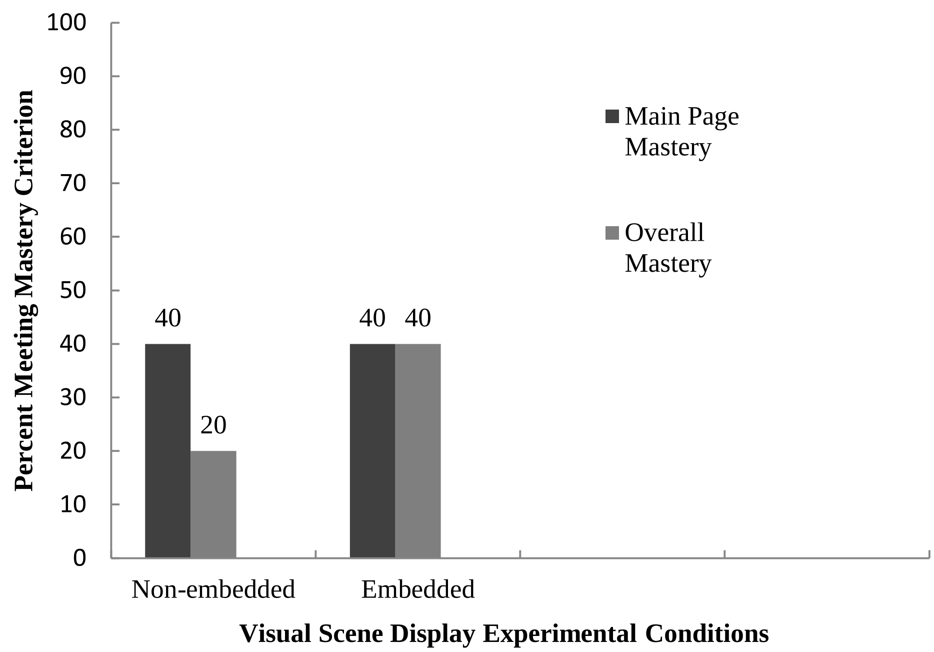


Figure 11. Comparison of main page mastery versus overall mastery between experimental conditions.



References

Allen, H. (1957). A new picture series for preschool vision testing. *American Journal of Ophthalmology, 44*, 38–41.

American National Standards Institute (ANSI). 1989. *American national standard specifications for audiometers*. ANSI S3.6-1989. New York: ANSI.

Beukelman, D. R., & Mirenda, P. (2005). *Augmentative and alternative communication: Supporting children and adults with complex communication needs* (3rd ed.). Baltimore, MD: Paul H. Brookes.

Brownell, R. (2000). *Receptive One-Word Picture Vocabulary Test, Second Edition* (ROWPVT). Novato, CA: Academic Therapy Publications.

Campbell, R., Baron-Cohen, S., Walker, J., (1995). Do people with autism show a whole face advantage in recognition of familiar faces and their parts? A test of central coherence theory. Unpublished manuscript, Goldsmiths College, University of London.

DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science, 238*, 1556-1557.

DeLoache, J. S. (1995a). Early symbolic understanding and use. In D. Medin (Ed.), The psychology of learning and motivation (Vol. 33, pp. 65–114). New York, NY: Academic Press.

DeLoache, J. S. (1995b). Early understanding and use of symbols. *Current Directions in Psychological Science, 4*, 109–113.

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences, 8*, 66–70.

DeLoache, J. S., & Burns, N. M. (1994). Symbolic functioning in young children. *Journal of Applied Developmental Psychology, 15*, 513–527.

Deruelle, C., Rondon, C., Gepner, B., Tardif, C., (2004). Spatial frequency and face processing in children with autism and Asperger syndrome, *Journal of Autism and Developmental Disorders, 34*, 199-210.

Dixon, L. S. (1981). A functional analysis of photo-object matching skills of severely retarded adolescents. *Journal of Applied Behavior Analysis, 14*, 465-478.

Drager, K., Light, J., Carlson, R., D'Silva, K., Larsson, B., Pitkin, L., & Stopper, G. (2004). Learning of dynamic display AAC technologies by typically developing 3-year-olds: Effects of different layouts and menu approaches. *Journal of Speech, Language, and Hearing Research, 47*, 1133–1148.

Drager, K., Light, J., Speltz, K., Fallon, L., & Jeffries, K. (2003). The performance of typically developing 2 1/2-year-olds on dynamic display AAC technologies with different system layouts and language organizations. *Journal of Speech, Language, and Hearing Research, 46*, 298–312.

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test-III*. Circle Pines, MN: AGS.

Faraway, J., (2005). Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models (texts in statistical science). Boca Routon, FL: Chapman and Hall/CRC.

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (1993). *The MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual*. Baltimore: Paul H. Brookes Publishing Co.

Fuller, D. R., & Lloyd, L. L. (1991). Toward a common usage of iconicity terminology. *Augmentative and Alternative Communication, 7*, 215–219.

Johnston, S., Reichle, J., Feeley, K., & Jones, E. (2012). *AAC Strategies for individuals with moderate to severe disabilities*. Baltimore, MD: Paul H. Brookes Publishing Co., Inc.

Light, J. (1997). “Let's go star fishing”: Reflections on the contexts of language learning for children who use aided AAC. *Augmentative and Alternative Communication, 13*, 158–171.

Light, J., & Drager, K. (2007). AAC technologies for young children with complex communication needs: State of the science and future research directions. *Augmentative and Alternative Communication, 23*, 204–216.

Light, J., Drager, K., McCarthy, J., Mellott, S., Millar, D., Parrish, C., Welliver, M. (2004). Performance of typically developing four- and five-year-old children with AAC systems using difference language organization techniques. *Augmentative and Alternative Communication, 20*, 63–88.

Lloyd, L. L., & Fuller, D. R. (1990). The role of iconicity in augmentative and alternative communication-symbol learning. In W. I. Fraser (Ed.), *Key issues in mental retardation research* (pp. 295–306). London, UK: Routledge.

Lund, S., Millar, D., Herman, M., Hinds, A., & Light, J. (1998). Children's pictorial representations of early emerging concepts: Implications for AAC. Paper presented at the American Speech Language Hearing Association Annual Convention, San Antonio, TX.

Mirenda, P., & Locke, P. (1989). A comparison of symbol transparency in nonspeaking persons with intellectual disability. *Journal of Speech and Hearing Disorders, 54*, 131–140.

Mizuko, M. (1987). Transparency and ease of learning of symbols represented by Blissymbols, PCS, and Picsyms. *Augmentative and Alternative Communication, 3*, 129–136.

Mizuko, M., & Reichle, J. (1989). Transparency and recall of symbols among intellectual handicapped adults. *Journal of Speech and Hearing Disorders, 54*, 627–633.

Mizuko, M., Reichle, J., Ratcliffe, A., & Esser, J. (1994). Effects of selection techniques and array sized on short-term visual memory. *Augmentative and Alternative Communication, 10*, 237-244.

Namy, L., Campbell, A., & Tomasello M. (2004): The Changing Role of Iconicity in Non-Verbal Symbol Learning: A U-Shaped Trajectory in the Acquisition of Arbitrary Gestures, *Journal of Cognition and Development, 5,* 37-57.

Olin, A., Reichle, J., Johnson, L., Monn, E. (2010). Examining Dynamic Visual Scene Displays: Implications for Arranging and Teaching Symbol Selection. *American Journal of Speech-Language Pathology, 19*, 284-297.

Patefield, W. M. (1981) Algorithm AS159. An efficient method of generating r x c tables with given row and column totals. *Applied Statistics, 30*, 91–97.

Paul, R. (2001). *Language disorders: From infancy through adolescence.* Philadelphia: Mosby.

R Development Core Team (2011). R: A language and environment for statistical. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Ring H., Baren-Cohen S., Wheelwright S. (1999). Cerebral correlates of preserved cognitive skills in autism: A functional MRI study of embedded figures task performance, *Brain*, *122*, 1305-1315.

Stephenson, J. (2009). Iconicity in the development of picture skills: Typical development and implications for individuals with severe intellectual disabilities. *Augmentative and Alternative Communication, 25*, 187-201.

Shah A., Frith U., (1983). An islet of ability in autism: a research note. *Journal of Child Psychology and Psychiatry 24*, 613-620.

Thistle, J., Wilkinson, K., (2009). The effects of color cues on typically developing preschoolers' speed of locating a target line drawing: Implications for augmentative and alternative communication display design. *American Journal of Speech-Language Pathology*, *18,* 231-240.

Tomasello, M., Striano, T., & Rochat, P. (1999). Do young children use objects as symbols? *British Journal of Developmental Psychology, 17,* 563–584.

Wilkinson, K., & Snell, J. (2011). Facilitating children's ability to distinguish symbols for emotions: The effects of background color cues and spatial arrangement of symbols on accuracy and speed of search. *American Journal of Speech-Language Pathology, 20*, 288-301.

Yates, F. (1934). Contingency tables involving small numbers of the χ2 test. *Journal of the Royal Statistical Society Supplement, 1*, 217–235