Preliminary Exploration of the Effect of Background Color on the Speed and Accuracy of Search for an Aided Symbol Target by Typically Developing Preschoolers

Krista M. Wilkinson, Ph.D.

Communication Sciences and Disorders
Pennsylvania State University and Shriver Center of the University of Massachusetts Medical School

Bridgett Coombs

Communication Sciences and Disorders, The Pennsylvania State University University Park, Pennsylvania

Aided augmentative and alternative communication can be used successfully with individuals with communication disabilities. Recent studies suggest that, where possible, arranging symbols based on internal color (placing red fruits together) facilitates search for a target symbol by children with and without Down syndrome (Wilkinson, Carlin, & Thistle, 2008). We explored whether color cuing of symbol background might offer similar benefits for symbols that cannot readily be arranged by internal color. Ten nondisabled preschoolers engaged in computer search tasks for line drawings representing common animals when the line drawings appeared on white backgrounds compared to color-saturated backgrounds that cued the subcategory to which the target belonged (land mammal, sea creature, bird, insect). Older children showed no consistent enhance-
Introduction

According to Beukelman and Mirenda (2005), 3.5 million Americans (1.3% of all individuals) have a communication disability that limits their access to natural speech as a primary means of communication. Causes of communication disorders can include intellectual disability, autism, cerebral palsy, apraxia of speech, stroke, traumatic brain injury, spinal cord injury, and so forth. Limitations in speech and communication can severely reduce quality of life.

Augmentative and alternative communication (AAC) is one means by which individuals with complex communication needs can improve their communicative functioning. AAC is defined by the American Speech-Language-Hearing Association (ASHA) as the “attempts to study and when necessary compensate for temporary or permanent impairments, activity limitations, and participation restrictions of persons with severe disorders of speech-language production and/or comprehension, including spoken and written modes of communication” (ASHA, 2005). An AAC device can be electronic or nonelectronic and may consist of words, pictures, symbols, and so forth. AAC can be used successfully with individuals of all ages with diverse communication needs (as cited in Beukelman & Mirenda, 2005, p. 8).

One limitation to using AAC is the rate and accuracy in which users can communicate their message. Foulds (1987) found that AAC users communicated at a rate of 15 words or less per minute, a rate that is fractional compared to natural speakers (as cited in Beukelman & Mirenda, 2005, p. 67). One reason that the rate of communication is slow for AAC users compared to natural speakers is because they must choose the item (e.g., letter, word, symbol, or picture) they wish to communicate one item at a time from an external physical display.

Slower communication rates may also relate to how readily the individual can detect, perceive, and identify the AAC symbols that are placed on a display. Wilkinson, Light, and colleagues (Light, Wilkinson, & Drager, submitted; Wilkinson & Jagaroo, 2004) have argued that the organization of symbols on an AAC device, such as symbol location, size, and color, may be very important for increasing the efficiency and effectiveness of communication. Such display considerations interact with individual characteristics of the AAC user (such as his or her cognitive, language, sensory, and motor abilities) and the task demands to influence functional communication outcomes (Light et al., submitted).

One dimension of display construction that has received substantial clinical attention as well as recent research concerns color. In basic research from visual cognitive science, color has been determined to be a powerful influence over a number of behavioral measures, such as detection, identification, or classification of a stimulus (see Wilkinson & Jagaroo, 2004, for a review related to AAC). Clinically, commercially available symbol sets allow clinicians to manipulate both the internal color of a symbol, where appropriate (coloring an apple red or green, or leaving it white inside) as well as its background. Use of color cues at a general level
conceivably might enhance the system’s appeal (Light, Drager, & Nemser, 2004), and clinical recommendations have suggested the use of color-coded backgrounds for symbols as a means to cue word-class category membership (objects, action labels; Goossens’ et al., 1999).

Despite its promise and widespread availability, the use of color cuing has only recently received direct research attention within AAC. The majority of studies concern the use of symbol-internal color cuing (coloring an apple red, or a banana yellow) as a means to facilitate rapid symbol detection and/or identification. For instance, Wilkinson, Carlin, and Jagaroo (2006) explored how readily typically developing preschool children could find a target symbol in an array of 8, when the symbols either all shared an internal color (8 red symbols), half the symbols shared colors (4 red and 4 yellow), or all 8 symbols were unique colors. Responses were faster and more accurate when the symbols had unique internal colors or when only a subset shared colors than when they were all red.

In a follow-up study, Wilkinson, Carlin, and Thistle (2008) examined shared symbol-internal color more closely, asking: When some symbols on a display share a color (there is more than one red item, for instance), does it matter how they are arranged? They sought to determine if it is better to cluster symbols with the same internal color or to distribute them throughout the array. Participants included 8 typically developing children over the age of four, 8 typically developing children under the age of four, and 10 children with Down syndrome. Participants were asked to locate target symbols for common fruits and vegetables under two conditions. In one condition, symbols with the same internal color were clustered together to create small, 4-symbol miniclusters of like-colored symbols. In the second condition, symbols with the same internal color were distributed throughout the array. When symbols that shared internal color were clustered together, speed in locating a target was increased by all participants, and accuracy was increased in the younger preschoolers and in participants with Down syndrome.

This research suggests that symbol internal color may exert an influence over responding, with particular facilitation when the like-colored symbols are clustered together. However, in the clinical literature the recommendations most typically concern not symbol-internal color cues, but rather color cues in the symbol background (e.g., Goossens’ et al., 1999). Moreover, preliminary descriptive work has suggested that when constructing displays clinicians may be more attuned to use of background color cuing than symbol internal color cuing (McFadd & Wilkinson, 2010). Wilkinson and colleagues (2008) in fact have speculated that background color cuing might serve the same function of facilitating search for a target, for symbols whose internal color cannot be manipulated (for instance, animals, or faces, where the color is often constrained). Yet, little research has examined the actual role of background color on search. These recommendations and speculations remain untested, empirically.

In one of the only related studies, Thistle and Wilkinson (2009) began to examine the role of symbol internal color as compared to symbol background color. Thirty-eight typically developing children between the ages 2:10 and 5:4 participated. Four experimental conditions were given: line drawings with relevant symbol internal color and no background color (a red strawberry on a white background), line drawings with no symbol internal color but a related background color (a white strawberry on a red background), line drawings with no symbol internal color and an unrelated background color (a white strawberry on a blue background), and line drawings with a relevant symbol internal color placed on an irrelevant background color (a red strawberry on a blue background). In younger children, as long as there was symbol internal color cuing, speed of locating the target was enhanced. Use of background color cuing in the absence of symbol internal color was detrimental to search performance, how-
ever. Although this effect was of statistical significance in younger children, it was not present in older ones.

The present research is a further preliminary exploration of the role of background color cuing. Wilkinson et al.’s (2008) study of symbol-internal color cuing had suggested that clustering symbols that shared internal colors facilitates search. Thistle and Wilkinson’s study (2009) suggested that background color is not detrimental to performance, if internal color is also present. However, in that study the presence of the symbol internal color cue was confounded/correlated with the background cue, because the symbols that appeared on a certain colored background all shared a symbol internal color as well (thus, all red fruits/vegetables appeared on blue backgrounds, yellow fruits/vegetables appeared on pink backgrounds, green fruits/vegetables appeared on orange background). What if there was symbol internal color, but it did not offer such a reliable cue, for instance, when the symbol internal color cannot be manipulated, reasonably changed, or clustered together? We expected that the color cuing of the symbol background would indeed aid in accuracy and speed of symbol selection despite the absence of symbol internal color cueing.

**Methods**

**Participants**

Although 13 participants were consented for the study, three children were excluded. One was excluded due to performance outside of two standard deviations of the norm on a standardized vocabulary test (see below) and two were excluded because their data showed substantial variations such that the standard deviation of the reaction times across repeated experimental sessions was greater than 33% of the mean reaction time (indicating wide swings in the session to session performance).

Typically developing children are commonly used in AAC research as a means to determine the expected cognitive levels of children and to limit other factors, such as physical and intellectual disabilities (Higginbotham, 1995). Participants aged three and older were chosen because children demonstrably understand that symbols represent real objects on a wide variety of measures (Beukelman & Mirenda, 2005; DeLoache, 1995, 2005) and because previous work had indicated that they would be capable of manipulating a mouse as an access mode (e.g., Thistle & Wilkinson, 2009; Wilkinson et al., 2008).

Two age groups were studied. The younger group included children under the age of 48 months (mean chronologic age = 44.8 months, range = 42–47 months). The older group included children over the age of 60 months (mean chronologic age = 62.4 months, range = 60–66 months). Each group contained five participants. The children attended university affiliated preschool and daycare classrooms in central Pennsylvania. Two of the 10 final participants represented diverse ethnic and cultural populations.

A standardized test of receptive vocabulary, the Peabody-Picture Vocabulary Test-IV (PPVT-IV) was given. For the younger group, the mean standard score on the PPVT-IV was 104.4 (range = 93–121). For the older group, the mean standard score on the PPVT-IV was 109 (range = 97–118). All participants whose data were used had mean standard scores or 95% confidence intervals within expected age ranges on the PPVT-IV.

**Materials and Displays**

All procedures were run via computer using a software program developed for this study (adapted from Dube, 1991). Presentation of the stimuli and recording of participant responses were controlled entirely by the program. Children accessed the computer using a mouse. Depending on the location, the computer used was a 20” iMac or a 17” laptop Macintosh. Although the comput-
ers were slightly different from one another, the software program ensured that the size of the symbols and the displays were identical across the two computers.

Stimulus materials consisted of digital photos obtained from the internet and line-drawing symbols obtained from the software program, Boardmaker™ (Mayer-Johnson, Inc.). Stimuli represented common exemplars from four general classes of animals; land mammals (dog, goat), sea creatures (shark, octopus), birds (parrot, blue-bird), and bugs (wasp, fly). These stimuli were chosen because they are likely to be highly recognizable to children at this age.

### General Procedures

Research sessions took place either in the preschool classroom or in private room in the preschool building. When the computer was used in the classroom, it was set up next to the other computers that already existed in the classroom and the children’s backs were faced away from the other children in the classroom to minimize distraction by the other classroom activities.

The research task consisted of 0-delay matching to sample in which the sample appeared in the middle of the screen. Upon mouseclick to the sample, it disappeared and was replaced by the choice array. Reinforcement for correct responses was provided by the program in the form of a herald sound. Silence followed each incorrect response. After each session, the participants were rewarded a sticker of their choice as reinforcement before returning to their class.

### Preassessment

Preassessment was conducted before the actual data collection to familiarize children with the basic task (finding the match) and with the stimuli that they would see in the experiment. In the pre-assessment, the sample consisted of a line-drawing and the choices consisted of the four photographs that represented animals from the same group (mammals, sea creatures, birds, or bugs). The four choices were presented in the four corners of the computer screen. All of the line drawings were presented with a white background during the pre-assessment. Participants performed at very high accuracies on preassessment (mean accuracy = 94% for younger participants, 92% for older participants).

### Experimental Conditions and Procedures

In the experimental conditions, the sample consisted of the photograph and the choice array consisted of the line-drawings (the reverse of the preassessment; reversing the roles of the sample and choices was conducted to reduce any carryover bias from preassessment to experimental sessions). Each child underwent study in two conditions, one in which the line drawings appeared on a white background (“white” experimental condition) and the other in which the line drawings appeared on a background saturated with a color (“color” experimental condition). The conditions were run on separate days. Six participants underwent the white background condition first, and the other four underwent the color condition first.

In each experimental condition, 16 symbols were presented that represented four exemplars from the four animal types (land mammals, sea creatures, birds, bugs). Based on the findings from Wilkinson et al. (2008) that suggested clustering is the optimal display organization, the four exemplars within each animal group were placed together in quadrants of the display. Because participants underwent both experimental conditions, we did not want the same set of exemplars to be used in both conditions. Consequently, different arrays that contained comparable animal classes but different exemplar animals were used in each condition. Figure 1 illustrates the displays, including: (a) the presence/absence of back-
ground color, (b) the clustering of related exemplars by quadrant, and (c) the different exemplars used under the two conditions.

One caveat is necessary. Of the 10 participants, 7 underwent study with the primary colors illustrated in Figure 1. Three, however, underwent study with the background color being a more neutral version of the colors from Figure 1. The reason for this is because we were interested in exploring whether responding might be different under neutral as compared to primary color background conditions. The influence of this alteration is considered in the results and discussion.

Dependent Measures

The dependent measures were the overall accuracy in the session and the mean reaction time to find the target after the onset of the choice array. Accuracy was calculated as the number of correct selections across each 16-trial session. Consistent with previous research (Wilkinson et al., 2006) reaction time was calculated as the median speed of selection for correct choices only. Median was used in order to minimize the influence of trials in which excessively long response times occurred, as it is less sensitive to outliers than the mean.

Results

In terms of accuracy, participants in the younger group found the target symbol with greater accuracy when the background was white (95% correct) than when the background had color (86% correct). Older participants showed a similar but far less pronounced pattern (mean for white = 91%; mean for color = 89%). Paired t-tests conducted separately for the data on younger children and for older children indicate that these differences were not of statistical significance for either group.

In reaction time, younger participants demonstrated enhanced performance when the background was white over when it had a color cue, while older participants showed little to no such effect. The median reaction times averaged across participants in each of the two groups are illustrated in Figure 2. Paired t-tests evaluated whether the difference between white and color background conditions was of statistical significance, for each group (with Bonferroni correction for the two tests). For the younger children, reaction time was significantly slower when color cues were present than with white backgrounds ($t(4) = 7.55, p = .002$); no significant effect was found for the older group.

Figure 1. Examples of displays under the two experimental conditions.
In exploratory work such as this, when there are few participants contributing to analysis, it is prudent to evaluate the extent to which group means represent the pattern of performance across individual participants. We therefore plotted in Figure 3 the data for each individual participant to evaluate the extent to which the mean represented the individuals. For ease of viewing, we plotted the difference between reaction time in the color cue condition versus the white background condition. In other words, we subtracted the median reaction time for each participant under the white background condition from the median reaction time under color background condition. Bars that are positive represent participants who showed enhancement under white background; bars that are negative represent enhancement by color background. Also noted in Figure 3 is whether the participant underwent study with primary color backgrounds (P) or neutral color backgrounds (N).

Figure 3 illustrates that all five younger participants performed better when the symbols had a white background rather than a colored background, and this advantage was greater than 1 second for all participants (mean = 1.3 seconds). In contrast, whereas two of the older participants showed a similar pattern of over 1 second advantage in the white background condition, two showed little difference and one showed an apparent advantage in the color condition. Thus, the individual data patterns reflect the group-level results of the two t-tests conducted (a consistent effect for younger participants, no consistent effect for older ones).

**Discussion**

The research sought to explore the proposition that the use of a background color might facilitate the speed and accuracy of locating a target line drawing in an array of symbols when internal color cannot be changed. We had hypothesized that the addition of background color would serve to
aid children in finding symbols, by highlighting subsets of items within which to search. In fact, these preliminary results appear to be quite the opposite. Although older children showed a mixed pattern of responding, all five of the younger children showed a consistent impediment to speed of search when color cues were added to the background, as compared to when the background was left white. This suggests that for younger children, use of background color cues may in fact be detrimental to facilitating rapid symbol detection.

**Implications: Symbol Internal Versus Background Color Cuing**

These results, although contrary to our initial expectations, were not actually inconsistent with prior findings by Thistle and Wilkinson (2009) concerning symbol background cuing. Specifically, Thistle and Wilkinson (2009) studied typically developing children and found that when background color is added to symbol arrays in which symbols with similar internal colors are clustered (red fruits/vegetables appear on blue backgrounds, yellow fruits/vegetables appear on pink backgrounds, green fruits/vegetables appear on orange background), search was not adversely affected. When the symbol internal color was removed and only the background color remained (white fruits appeared on the clustered blue, pink, and orange background colors), however, the performance of younger children dropped significantly. Older children showed little change in performance.

**Figure 3.** Difference between reaction times when symbols have color-saturated backgrounds as compared to white backgrounds, represented in a difference score (color-white) for each of the 10 participants (positive bars reflect advantage under white background condition, negative bars reflect advantage under color background condition).
Thistle and Wilkinson’s (2009) strategy of removing the symbol internal color resembles the conditions in our study, insofar as the symbol internal color was not available as a cue. Although the symbols in our study did have internal colors, those colors present within our symbols could not be manipulated; there are only so many colors a dog or a cat can be. In both the current study and that of Thistle and Wilkinson (2009), absence of the symbol-internal color cue had a detrimental effect on efficiency of search for younger, but not older children. Consequently, it would seem that the function of color cuing the internal portions of a symbol may be very different from the potential function(s) of symbol background coloring.

Age-Related Differences and Display Construction

This research shows that it is important for clinicians and educators to take into account the age of the child they are working with to determine the fastest, most accurate way for the child to select symbols when using their device. Indeed, the age-related difference in performance was not wholly surprising, given the consistent findings in previous research that younger children may be more dependent on physical/perceptual cues than older ones (Thistle & Wilkinson, 2009; Wilkinson et al., 2006, 2008). However, the fact that older children in both Thistle and Wilkinson (2009) and the current research did not show the detriment in performance when background color cuing was present may be very important.

In particular, this age-related finding implies once more that the dependence on the physical characteristics of the display may be lessened with maturation/experience, such that aided AAC displays for older or more advanced communicators may not require the same level of control over the physical features as for chronologically or developmentally younger children. This may be quite important because the use of background color cuing in the current recommendations is often times for higher order language functions, such as word-class category (e.g., objects are placed onto one background color, action labels onto another; Goossens’ et al., 1999) that would be most likely to be implemented for older or more advanced communicators. Indeed, we did not test the role of color in word class category cuing, and our results therefore cannot be interpreted in relation to that type of cuing nor as they might apply to individuals for whom syntactic development is a current goal. It may be that indeed, the function of background color is nonlinear, with the background color cuing offering benefits when tested with older children and/or advanced categorization. Clearly, this is an important area for further study.

Limitations and Future Research

This study was not intended to serve as the final answer concerning background color cuing, and the results must be interpreted with caution until further research is conducted. Indeed, it seems likely that there will be many circumstances under which background color cuing is facilitative, and it will be important to delineate those. We discuss some of the limitations of the current research and outline some potentially fruitful avenues for future research.

Sample Size

Clearly, a major limitation of this study is the number of participants. Each of the two groups contained only five children, and there is always the possibility that the findings are not representative. Although the consistency of the findings within the younger group (illustrated in Figure 3) suggests that the findings may be robust, clearly, there is a need for future studies that contain more participants.
Developmental Implications

There are several important gaps remaining to be studied concerning the developmental nature of the work thus far. First, it would be interesting to examine performance in children aged 48 to 59 months. Such children would be in between the groups analyzed in this study in regard to age. Analyzing a “middle” group may give us more information on the how the speed and accuracy of symbol selection change as children grow and develop. Specifically, we found that children under age 4 were distracted by the color background, whereas those over age 5 showed no effect on performances with or without the background. Is this shift linear, in which the interference from the background color gradually decreases, or is it instead something more qualitative, in which children are either significantly affected or not affected at all? If it is qualitative, when does this shift occur?

Another important piece of information concerns whether there is a point, developmentally or functionally, at which color background begins to facilitate search. Only one of the five older children showed a superior performance under the color-background condition. If these preliminary data are upheld in further study, this suggests that the color-coded background cuing, although not having the same detrimental effect on search that it had for younger children, also had no facilitative effect for children between 5 and 6 years of age. Clearly, there are other possible positive outcomes that might be facilitated by background color cues, particularly over the longer term. Perhaps color-coded background has positive effects on speed of combining more than one symbol together, as compared to finding a single symbol as examined in this research. It also is possible that color-coded backgrounds aid in acquisition over time, rather than the initial search on a relatively unfamiliar display. Finally, it is possible that color-coded backgrounds facilitate search in individuals older than the ones studied here, that is, over the age of 6 years. Clearly, studies of the effect of color-coded backgrounds on other dependent measures and with participants who are older than those studied here will be of great importance.

Finally, the sample was also composed solely of children without disabilities. Although initial findings would suggest that patterns of performance in children with Down syndrome may resemble those of their nondisabled peers (Wilkinson et al., 2008), this may not be the case for those with autism or other disorders (Wilkinson, Coombs, & McIlvane, 2009). Research is desperately needed to extend these findings to groups of children who have communication disabilities, who would be likely users of aided AAC.

How Clinically Significant are the Improvements in Reaction Time?

In research like this, a critically important question concerns whether an improvement that reaches statistical significance is of any relevance clinically. In this case, the younger children showed a detriment of about 1.3 seconds when background color cuing was present. Is this a substantial enough interference to warrant removal of color cues from their displays?

As noted in the introduction, on average AAC users who use direct selection communicate at a rate of 15 words or less per minute. This rate is fractional compared to that of natural speakers (Foulds, 1987). The present research suggests that the rate of communication of younger children will be faster if background color is not included. Although a single symbol selection would not necessarily be perceptually distinguishable, if symbol rate is slower by a second and the user seeks to communicate about three, or four, or five symbols, then clearly the detriment to rate of message preparation will become obvious to listeners (being up to 5 seconds slower) and likely indeed alter the communication patterns.
We would never suggest that by itself the use or nonuse of color cuing is the only consideration in constructing aided AAC. Nor would we suggest that altering color cues will make the difference between success and failure with an aided system. However, manipulation of color cuing is fairly readily accomplished. We would suggest, therefore, that consideration of color cuing is important as one of many tools the clinician can use to facilitate communication. The current research would suggest that for younger children, white backgrounds offer better support for search than colored ones.

Use of Taxonomic Categories

A third limitation of this study is the use of the subordinate categories presented in each grid to cluster the AAC symbols according to the type of animal it represents (e.g., “mammal,” “bird,” “fish,” or “insect”). According to Fallon, Light, and Achenbach (2003), typically developing children have difficulty using this type of organization until ages 6 or 7. They suggest that AAC symbols should not be organized by the category in which they belong until the AAC user is developmentally older than age 6 or 7. Because all of the participants of this study were under the age of 66 months, the clustering of symbols according to categories may have affected the accuracy and latency of the participants’ responses because they have not developed the skill of differentiating among different categories. Although this issue would not bear on the differences between the background color conditions, future research is needed to determine if there is a difference in the speed and accuracy of symbol selection when the symbols are scattered throughout the grid instead of clustered by category.

Use of Saturated Symbol Backgrounds

In this research, we chose to color the entire background of the symbol. At a descriptive level, it seems not to have affected outcomes when a primary versus a neutral shade was used in the saturation, although clearly this variable was not well controlled. An alternative is to present the symbol on a border (a red border around the white symbol). This alternative would retain the black-white contrast of the symbol and its background, but still offer a potential color cue. Would this have helped the younger children in this study? Another recent study of color cuing of symbols representing emotions (Snell & Wilkinson, 2009) offers preliminary data to suggest that border color cues are equally problematic, however, the issue is far from being resolved. Yet a third alternative may be to place symbols on color strips, rather than have the color attached to the individual symbols. Research is currently underway to explore these possibilities.

Influence of Array Size

In the current study, participants were instructed to match a digital picture of an animal to its corresponding line drawing symbol. The symbols were placed in an array of 16 line drawing symbols on each grid in the computer program. Thistle and Wilkinson (2009) used an array of 12 line drawing symbols in each of their grids. It would be interesting to learn what the results of this study would be when the demand is higher or lower. For example, would the latency and accuracy results of typically developing preschoolers be different if the array was smaller (e.g., an array of 6 or 8) or larger (e.g., an array of 24, 48)? Future research is needed to determine an ideal number of symbols contained in an array.

Conclusions

The results of this study are necessarily preliminary, given the sample size, composition, and other limitations identified above. Nonetheless, the data appear to be converging on the conclusion that color coding of the background of the symbol, although
not affecting older children’s search performance, may in fact be detrimental to younger children. In that event, the use of color coding may be better reserved for the types of language goals and displays more often used by older children, who may indeed benefit from this perceptual cue that may be distracting for younger ones. Clearly, further research is necessary.

Acknowledgments: This study was conducted as part of a master’s project by the second author in the department of Communication Sciences and Disorders of the Pennsylvania State University, under the supervision of the first author. The research was supported by P01 HD25995 of the University of Massachusetts Medical School-Shriver Center and the Pennsylvania State University department of Communication Sciences and Disorders. Many thanks to the staff and families of the Bennett Center and the Child Development Laboratories of Pennsylvania State University.

Address Correspondence to: Krista Wilkinson, Department of Communication Sciences and Disorders, Pennsylvania State University, 404H Ford Building, University Park, PA 16802; Tel: (814) 863-2206; E-mail: kmw22@psu.edu

References


