Published in final edited form as:

Augment Altern Commun. 2012 September ; 28(3): 137–147. doi:10.3109/07434618.2012.704522.

Considerations for the Composition of Visual Scene Displays: Potential Contributions of Information from Visual and Cognitive Sciences (Forum Note)

Krista M. Wilkinson, Janice Light, and Kathryn Drager

The Pennsylvania State University

Abstract

Aided augmentative and alternative (AAC) interventions have been demonstrated to facilitate a variety of communication outcomes in persons with intellectual disabilities. Most aided AAC systems rely on a visual modality. When the medium for communication is visual, it seems likely that the effectiveness of intervention depends in part on the effectiveness and efficiency with which the information presented in the display can be perceived, identified, and extracted by communicators and their partners. Understanding of visual-cognitive processing – that is, how a user attends, perceives, and makes sense of the visual information on the display – therefore seems critical to designing effective aided AAC interventions. In this Forum Note, we discuss characteristics of one particular type of aided AAC display, that is, Visual Scene Displays (VSDs) as they may relate to user visual and cognitive processing. We consider three specific ways in which bodies of knowledge drawn from the visual cognitive sciences may be relevant to the composition of VSDs, with the understanding the direct research with children with complex communication needs is necessary to verify or refute our speculations.

Aided augmentative and alternative (AAC) interventions have been demonstrated to facilitate a variety of communication outcomes in persons with intellectual disabilities, including self-expression and, for many individuals, understanding of language input (e.g., Bondy & Frost, 2001; Bopp, Brown, & Mirenda, 2004; Harris & Reichle, 2004; Johnston & Reichle, 1993; Ronski & Sevcik, 1996; see Mirenda & Iacono, 2009, or Wilkinson & Hennig, 2008, for reviews). While aided AAC has enriched the lives of many individuals, there are still individuals who either resist adopting the systems or abandon the systems after initial efforts (Culp, Ambrosi, Berninger, & Mitchell, 1986; Phillips & Zhao, 1993). To understand this phenomenon, Johnson et al (2006) conducted factor analysis on surveys from 275 clinicians concerning the features reported to contribute to success versus abandonment of AAC systems. Clinicians reported that success was predicted by “system characteristics and fit,” while abandonment was predicted by a “poor fit” between the features of a system and the characteristics of its user.

One system characteristic that may impact success is the composition of the AAC display. Most aided AAC systems rely on a visual modality. When the medium for communication is visual, it seems likely that the effectiveness of intervention will depend in part on the effectiveness and efficiency with which the information presented in the display can be perceived, identified, and extracted by the communicators and their partners (e.g., Wilkinson & Jagaroo, 2004). Understanding of visual-cognitive processing – that is, how a user attends, perceives, and makes sense of the visual information on the display – therefore seems critical to designing aided AAC interventions that are appropriate “fits” to individual clients. In this paper, we consider several characteristics of one type of aided AAC display,

that is, Visual Scene Displays (VSDs), in terms of their potential relationship to visual and cognitive processing.

Introduction to Visual Scene Displays and Goals of this Forum Note

In VSDs, representations of concepts are embedded in a full or integrated scene. This means that the entire image of the scene – whether a digital photograph, scanned image, or schematic line drawing – contains multiple concepts within it. An example of a VSD is presented in Figure 1, which depicts a young child's birthday; the VSD is realized through a digital photo of an adult holding the birthday cake and the child about to blow out the candles. Vocabulary associated with the event is programmed into the scene using aspects of the scene as hotspots to represent the language concepts. For instance, selection of the birthday cake might activate the underlying hotspot, resulting in the speech output “birthday cake” or singing of the birthday song; selection of the candles might retrieve the message “candles”; selection of the child's mouth might retrieve the message “blow” with a sound effect of blowing; selection of the child might retrieve the child's name, etc. It is important to understand that a VSD is not in any way restricted to communicating a single message or communicative intent. Rather, it is a milieu through which individuals might communicate multiple messages or intents. For instance, the birthday party scene might be used by the child to request to sing happy birthday, to comment on blowing out the candles, or to name the birthday girl.

Uses of VSDs and the Underlying Rationale

VSDs are currently used in many clinical settings, and have been integrated into technology offered by commercial manufacturers. Blackstone (2005) provided an in-depth review of some of the clinical applications of VSDs. As she noted, VSDs are typically digital photos or line drawings that can be highly personalized, contain extensive context to support understanding, and have the primary purposes of supporting conversation and social interaction, learning, and general communication activities like play, storytelling, and participation. They can be used with communicators with a variety of skill levels, when these communication purposes are necessary and for users who can benefit from the contextual support (often, beginning communicators such as very young children). The types of scenes included within VSDs can range from significant events, such as birthday parties and holiday celebrations, to daily activities such as playing on a playground or having a picnic.

The rationale underlying the use of VSDs with young children is that they tap into the rich, context-supported and event-based learning mechanisms that are thought to underlie language learning (Drager et al., 2003; Light et al., 2004; Wilkinson & Light, 2011). Specifically, Nelson (1986) and others (e.g., Bruner, 1983) have noted that children learn early language concepts within a rich, context-based environment in which diverse experiences contribute to the child's growing conceptual knowledge. Thus, children learn about and hear the label for dogs as they see dogs at the park, pat them at their relatives' houses, get kisses from them, and so forth. VSDs are intended to provide a similar context by embedding the language concepts within photographs of the events or routines familiar to the child, thus providing a visual form of this supportive context for the learning of concepts.

Current State of Knowledge about Visual Scene Displays

The research on VSDs is still in its infancy, although there has been a recent spate of studies examining the use of VSDs with young children - the focus of the current paper - as well as adults with acquired disorders (Drager, Light, Carlson, D'Silva, Larsson, Pitkin, et al., 2004;

Drager et al., 2003; Light et al., 2004; McKelvey, Hux, Dietz, & Beukelman, 2010; see Wilkinson & Jagaroo, 2004 for a discussion). Thus far, most of the research has involved VSDs realized through line drawings rather than photographic ones. In studies of young children, Drager and colleagues found that performance of typically developing 2 ½-year-olds in locating representations of language concepts was significantly better with VSDs than with traditional grid displays (Drager et al., 2003). Recent research by Light and Drager (2010) utilized a single subject multiple baseline design to investigate early AAC intervention using VSDs of familiar motivating events with nine infants and toddlers (ages 6-40 months at introduction) with complex communication needs (e.g., Down syndrome, cerebral palsy). From the very first session, all of the children took turns in social interactions using VSDs of familiar events realized through digital photographs and scanned images. Furthermore, the children demonstrated significant increases in their rates of turn taking as well as multiple other modes during intervention compared to baseline levels.

Goals of this Forum Note

VSDs seem to hold promise as a layout option for aided AAC systems for young children. The research has not, however, identified which elements of VSDs make for more or less effective displays, nor have the studies compared different compositions of VSDs to examine relative effectiveness. Thus, although suggestions have been offered concerning elements of VSDs that might support their successful use (Dietz et al., 2006; Light & Drager, 2008), there is currently little research examining those propositions directly.

Although little guidance exists within the field of AAC about visual cognitive processing and the composition of VSDs, the fields of visual cognitive science and neuroscience provide an emerging body of research focused precisely on the various factors that can influence visual cognitive processing of what are called “integrated” or “naturalistic” scenes.¹ In this paper, we seek to explore specific elements from this body of research to suggest avenues of inquiry that might be of potential relevance to researchers and clinicians who are interested in constructing photographic VSDs for children. We restrict our consideration to VSDs, rather than traditional grids. As Wilkinson and Jagaroo (2004) have noted, there are clear structural differences between the two displays. Processing of VSDs likely involves a different permutation of the visual processing neural network than grids, and as a result differ in the visual processing demands they impose on users. If so, principles that guide construction of one type of display may not transfer seamlessly to the other.

This effort is conducted in the spirit of previous interdisciplinary efforts of this sort, from one of the earliest calls to consider the cognitive sciences (Light & Lindsay, 1991) to recent examinations of the potential contributions of visual cognitive neurosciences for construction of aided AAC displays (Wilkinson & Jagaroo, 2004; Jagaroo & Wilkinson, 2008). Like the prior efforts, this review is not intended as a comprehensive analysis of the literature on visual scene processing. Rather, we seek to illustrate how specific research in that field might be relevant to the ways in which researchers and clinicians might approach

¹Any interdisciplinary effort such as this one requires the integration of disparate technical vocabularies. In cognitive neuroscience, the term for a stimulus (such as a photograph) consisting of an integrated scene is “natural/naturalistic scene.” While photographs of natural/naturalistic scenes are often the medium for the VSDs used in AAC, some photographs/natural scenes might make for effective VSDs, while others could be less effective. For example, some naturalistic scenes that might be studied in the visual cognitive sciences (e.g., photographs of a deserted beach, a distant mountain, or an empty room) might not serve as effective VSDs if they do not capture meaningful and motivating events that serve as the context to support a range of messages or intents by the communicator. It is therefore important to maintain the distinction between the idea of a photograph of a naturalistic scene, which can contain any integrated scene, and a photograph-based VSD that is intended to serve as a milieu to support the communication of specific functions. This paper seeks to integrate information about naturalistic scene perception for the purposes of understanding VSDs as a particular type of naturalistic scene (one that is intended to support communication). In this paper, we use the terminology that is consistent with the discipline in which the knowledge originated, but with the understanding that the extension is intended specifically for VSDs.

the design of VSDs. Readers are also cautioned that the considerations proposed in this paper, while generally very well-researched in visual cognitive neuroscience, have not yet been systematically investigated within the field of AAC. Future research is obviously required to test out these suggestions empirically with children with complex communication needs.

Structure of this Forum Note

In each section of this paper, we begin by identifying a clinical issue or problem in the design of VSDs. We then review the research from visual cognitive neuroscience that is related to that issue. Finally, we consider how this body of research might relate to or be useful when designing VSD displays, with an eye to acknowledging where possible differences might emerge when considering populations with disabilities.

We have selected three examples from the many possible topics: (a) the notion of the “complexity” of VSDs, (b) the role of humans in VSDs, and (c) the inclusion of events within VSDs. These examples reflect areas that have both immediate relevance for clinical VSDs as well as evidence from visual cognitive science and neurosciences research. We are not suggesting that these are the only aspects of VSDs that are important. Rather, these aspects are intended as examples of the issues involved in the design and use of VSDs for communication.

We have opted to put off consideration of one aspect of visual displays that is quite well-researched in AAC, that is, the issue of iconicity. This term refers to the ease with which the meaning of a representation or symbol may be guessed by a naïve viewer (e.g., Fuller, 1997), and there is a fairly thorough body of work that has compared symbols and symbol sets on the relative guessability, or iconicity, of single-meaning symbol representations. This research suggests that the greater the physical resemblance of a two-dimensional image to its referent, the more readily the symbol meaning is guessed; thus, in general, photographs of individual items (e.g., teacup) have been reported to be more iconic than line drawings of that same item, which in turn are more iconic than Blissymbols or arbitrary symbols like lexigrams (Bloomberg, Karlan, & Lloyd, 1990; Hurlbut, Iwata, & Green, 1982; Mirenda & Locke, 1989; Mizuko & Reichle, 1989). While we are fully convinced of the urgent need to address this issue, we have deferred its consideration for two reasons. First, the issues of iconicity within AAC are quite complex, and a thorough theoretical treatment is far beyond the scope of this paper. Just as important, the visual cognitive sciences literature does not appear to have a body of research on this issue from which to draw the kinds of parallels we deliberately seek to highlight here. A future, thorough consideration of iconicity and its implications for VSDs is clearly warranted.

The Notion of the “Complexity” of a Visual Scene Display

Historically, the field of AAC has focused on traditional grid displays with individual AAC symbols organized in rows and columns on the display. The “complexity” of the aided AAC stimulus has traditionally been defined in terms of the number of elements that make up an individual symbol or the number of symbols in the array or display on which the symbols appear. For instance, when examining the influence of complexity of symbols on learning outcomes, Fuller (1997) defined the term operationally as “the number of strokes required to produce a symbol. Symbols having one to five strokes were considered low in complexity while those with eight or more strokes were considered high” (p. 32). Similarly, Barton and colleagues (Barton, Sevcik, & Ronski, 2006, p. 15) equated the visual complexity of Bliss and lexigram symbols by matching the number of elements (geometric forms and/or strokes) within them. This measure has been reported to correlate with perceptions of complexity of Blissymbols (Fuller & Lloyd, 1987) and to influence learning outcomes in a paired associate

task, for nondisabled preschool children (Fuller, 1997). When applied to entire displays, the term complexity has been used in a similar manner; arrays with greater numbers of symbols have been considered more complex, in terms of demands on processing, than arrays with fewer symbols. Mizuko, Reichle, Ratcliff, and Esser (1994) reported that the number of symbols in an array influences short term memory for sequences of symbols in nondisabled preschoolers.

The “Complexity” of VSDs

When looking at the composition of VSDs, one immediate observation is that they typically contain numerous elements – some more relevant to the scene than others – and most of these elements are themselves made up of many lines, shapes, etc. For instance, the VSD of a child's birthday party illustrated in Figure 1 includes not only the key elements of the event - the child, the cake and the candle - but also other possibly irrelevant elements that contribute less to the general meaning of the VSD, like the high chair tray, the bib hanging off the chair, the doorknob, etc. If one considers the notion of “complexity” in terms of the sheer number of keystrokes or elements, then the VSD might be considered to be highly complex.

The numerous elements in the VSD might be argued to add to processing load of the display, thus adding significant demands on the user's perception and/or processing. If so, this added load might arguably offset any advantages offered by the visual and contextual support provided by the VSD. We turn now to a consideration of what the research in the visual cognitive sciences suggests about this issue.

Information about Related Issues in Visual and Cognitive Sciences

Within the last 10 years, a series of studies in the visual cognitive sciences have emerged that focus on processing of what are called “naturalistic scenes,” similar to those used in VSDs. This body of research has indicated that the way humans process naturalistic scenes is likely quite different from, and perhaps not predicted by, principles of visual attention derived from studies of simple stimuli such as letters or geometric shapes (e.g., see Braun, 2003, or Oliva & Torralba, 2007, for brief but informative reviews, or Kayser, Kording, & Konig, 2004, for consideration at the level of brain responses). In fact, the research in the visual cognitive sciences suggests that humans process naturalistic scenes very rapidly despite the presence of numerous elements (cf. Fletcher-Watson, Findlay, Leekam, & Benson, 2008). Humans recognize the overall context of naturalistic scenes as well as the constituent elements of these scenes (i.e., the people and objects depicted in the scene) within the very first look. In their comprehensive review of this literature, Oliva and Torralba (2007) noted that the speed of recognition of the overall context as well as the scene's constituent elements is accomplished within the very first glance at a speed of 200 milliseconds or less. Those authors noted: “A scene composed of contextually related objects is more than just the sum of the constituent objects,” and they went on to say that “Even if objects can be identified by intrinsic information, context can simplify the object discrimination by decreasing the number of object categories, scales, and positions that must be considered” (p. 525).

The behavioral and emerging neurological evidence suggest some reasons that these apparently complex scenes can be processed so quickly. At a general level, presence of visual context (such as that found in scenes) has been found to influence not only object recognition (see Oliva & Torralba, 2007) and patterns of attention in general (Braun, 2003; Mack & Eckstein, 2011), but also higher-order behaviors such as classification of stimuli (Fei-Fei, VanRullen, Koch, & Perona, 2005) and extraction of information about the scene (Venturino & Gagnon, 1992). Fei-Fei et al. (2005) demonstrated that the recognition of

object identity within scenes can be accomplished at very brief exposures, and even in the absence of directed attention; for instance, objects are recognized rapidly in scenes even when the participant is actually asked to do some other task that detracts attention from the actual scene. In contrast, performance classifying stimuli like geometric shapes, not presented within naturalistic scenes, is far worse than performance detecting objects that appear within naturalistic scenes, even though the arrays of shapes are seemingly “simpler” in terms of the number of elements. Recent brain imaging work has reportedly uncovered the cortical neural mechanisms associated with rapid natural scene categorization and object recognition (Peelen, Fei-Fei, & Kastner, 2009).

This capacity for rapid processing of scenes is perhaps not surprising since humans must regularly process this type of visual input in their everyday lives, beginning at infancy. One factor that may facilitate the rapid processing and recognition of scenes is that scenes exploit real-world experience and the recognition or activation of experience-based schemas (Braun, 2003; Oliva & Torralba, 2007; Venturino & Gagnon, 1992). Activation of such schemas allows prediction of elements that would be expected to occur within these scenes as well as the relative physical relationships between these elements (cf. Biederman, 1981).

One conclusion that may be drawn from this research is that an increase in the number of elements in a naturalistic scene may not necessarily require greater neural processing time, because the familiarity of the scene may trigger existing visual templates and/or schema that facilitate recognition and processing. This body of research suggests that within naturalistic scenes, the presence of context facilitates attention, detection, interpretation, and classification of scenes as well as their constituent parts. The added elements do not seem to add greater complexity from a processing standpoint; rather, they seem to facilitate processing.

How Might this Research Relate to VSDs and Individuals with Disabilities?

Like the natural scenes investigated in the visual cognitive science and neuroscience literature, VSDs like those in Figure 1 preserve the visual and conceptual relationships between the elements in the scene (i.e., objects, people) and the expected/familiar events, providing explicit cues to support recognition. For instance, in a VSD of a birthday party, the person represented would be expected to be carrying a cake toward the table, not carrying a dog or an umbrella or doing something else with the cake, such as putting it in the trash. Consequently, despite the potential increase in the number of constituent elements in VSDs, it seems quite possible that VSDs as a medium for AAC may exploit the same type of visual and neural processes that allow for rapid processing and interpretation of naturalistic scenes. If so, it could be possible that inclusion of meaningful elements in a VSD may actually facilitate processing because the presence of the features contributes to understanding the schema depicted.

The research from the visual and cognitive sciences focuses on participants (primarily adults) with intact visual systems. Questions remain when one considers how this information might extend to individuals with disabilities who might use AAC. Beginning communicators who might benefit from the use of VSDs to enhance their communication may be much younger and may demonstrate quantitative or qualitative differences in their processing. Because individuals with autism, for instance, have consistently been reported to demonstrate superiority in a range of visual search tasks (e.g., see Joseph, Brandon, Connolly, Wolfe, & Horowitz, 2009), they might be expected to perform particularly well at locating information within a VSD. It is also equally plausible to suggest the opposite, however - that individuals with autism would have more difficulty with VSDs due to their tendency to focus narrowly or in an overselective fashion on details but not on more global features of stimuli (Liss, Saulnier, Fein, & Kinsbourne, 2006; Mann & Walker, 2003;

Plaisted, O'Riordan, & Baron-Cohen, 1998; Plaisted, Swettenham, & Rees, 1999; Wilhelm & Lovaas, 1976) and/or because their visual attention to humans is mitigated relative to nondisabled individuals (Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Riby & Hancock, 2008, 2009). Finally, many individuals with complex communication needs are at risk for visual impairments, which might alter their ability to detect information within the VSD context. It seems fairly clear that a nuanced understanding of the speed with which humans can identify both the overall context of a scene (e.g., playing blocks in the living room, eating dinner in the kitchen), and also constituent elements within that scene (e.g., the blocks, the child, etc) will require direct investigation with individuals with complex communication needs.

Human Figures as Central Elements within VSDs

In addition to investigating the processing of VSDs as a whole, it is also important to consider the effects of specific constituent elements in VSDs (e.g., people, objects, events). Currently, some commercially-available VSDs consist of generic backdrops of locations that provide primarily object content (e.g., a line drawing depicting a kitchen with a fridge, stove, table, etc) and omit people within these scenes. If the rationale underlying VSDs is that they offer rich, context-based visual support based on experiences familiar to the child, based in the social communicative interactions within which communication occurs (e.g., Nelson, 1986), then it might be argued that a critical compositional element of VSDs is the inclusion of people engaged in events that support the communication messages and intents of the user.

Human Figures as Part of VSDs

Humans are an integral component of young children's experiences (e.g., playing ball with dad, blowing bubbles with mom). From birth, children are engaged in social interactions with partners that in turn foster communication and language functions. For example, early social transactions between infants and caregivers involve mutual smiling games that set the stage for the development of "proto-conversations" and later pragmatic development (Bruner, 1983; Wetherby & Prizant, 1993). VSDs incorporate social event experiences in order to support development and use of language concepts. It might therefore be argued that humans are an essential component of VSDs. We turn now to a review of visual attention to humans in scenes from the visual and cognitive sciences.

Information about Related Issues in Visual and Cognitive Sciences

Humans are compelling stimuli from a visual standpoint from early development. Within the first months of life, the most reliable stimuli to attract infant attention are faces, and in particular, human eyes (Gliga, Elsabbagh, Andravizou, & Johnson, 2009; Hopkins, 1980). Initially when the infant's visual system is still underdeveloped, even the presentation of very simple eyes elicits a response such as a smile; as babies grow and experience social interaction (and their vision matures), the power of animacy remains, but more of the face is required to elicit a smile (Bower, 1977; reviewed in Hopkins, 1980). The relation of this innate, or at the least early-emerging, attraction to humans may be of interest in the construction of VSDs.

The literature in cognitive sciences on behavioral responses to scenes with people as well as the literature in cognitive neuroscience on processing of human faces provides evidence that may be relevant to considering the critical salience or "social relevance" (Bracco & Chiorri, 2009) of humans in VSDs. One approach used to investigate the effect of humans in scenes involves a task that requires viewers to identify an element that changes between two versions of a scene (photograph) presented in rapid succession, but with a brief interval

between presentations (Bracco & Chiorri, 2009; New, Cosmides, & Tooby, 2007; Rensink, O'Regan, & Clark, 1997). This “change blindness” paradigm allows experimenters to examine the average time that it takes viewers to notice or find the changed element, with the assumption that the more salient or relevant an element is, the more rapid the search time will be. Bracco and Chiorri (2009) examined how readily 30 undergraduate students identified a change in a scene that either involved a person (e.g., a person at the periphery of the scene was absent or present) or an inanimate object (e.g., a branch on a tree overhanging a volleyball game was present or absent). Twenty-three photographs involving each type of change were created, holding constant the location, dimension, and saliency of the change. Participants were more accurate and faster detecting the change when the change involved a person target than an object target. A pattern of preferential attention to animate beings (including humans) was also reported in a similar change blindness procedure by New et al. (2007), even when controlling for low-level visual factors.

Further evidence for the salience of humans in scenes, even at the level of early visual processing, is offered by studies of attentional capture during scene viewing while eye-tracking equipment records exact point of gaze. For example, Fletcher-Watson et al. (2008) examined which of two different scenes presented in pairs attracted more rapid and longer attention when presented simultaneously, side by side. One of the scenes contained a human and the other did not. Forty such stimulus pairs (human present-human absent) were presented to 12 college-aged participants, and the location of the first look as well as the duration of time spent on each scene was recorded through the eyetracking technology. During free viewing, 11 of the 12 participants spent a significantly longer time viewing the scene containing the person than the one that did not in each trial. Furthermore, they also were significantly more likely to look first at the scene containing the person, and the first fixations tended to occur within the first 150 milliseconds after the stimuli were presented. In another study examining preferential looking patterns, Crouzet, Kirchner, and Thorpe (2010) presented images of faces side-by-side with other categories of stimuli such as vehicles, at very brief durations (400 milliseconds). Participants were instructed to look for one of the different types of images (e.g., face, vehicle). Even when participants were specifically instructed to look for non-face stimuli (e.g., vehicle), they were unable to keep from making quick saccades towards the faces prior to locating the non-face target. Finally, Smilek et al. (2006) evaluated eye gaze patterns of participants viewing two photographs, presented separately, one of a basketball game and one of an artist painting a model. Fixations in the scene were almost exclusively committed to the humans in the foreground, particularly to the face region. There were some fixations also to other meaningful elements (e.g., the basketball), but not to semantically less important ones.

Recent research by Wilkinson and Light (2011) has examined attention to human figures within photographs of scenes including humans that were small, offset, and/or presented with other objects in the scene that might compete for attention. The eye-gaze patterns of 19 college students were obtained while they viewed eight different photographs. In all photographs, a small human figure (occupying a mean of 6% of the photographic space) was presented alongside colorful or prominent distracters, such as a fully decorated Christmas tree, the China tower at Epcot™, or a large statue of soldiers marching. Even under these conditions, attention was drawn to the human in the scene more rapidly and for proportionally longer than virtually all other elements in the photographs. These data reinforce the power of human figures to attract visual attention within scenes irrespective of what else may be present.

How Might this Research Relate to VSDs and Individuals with Disabilities?

For children at the beginning stages of communication development, attracting and maintaining interest is likely a key step in establishing early AAC system use. It seems

reasonable to suggest that emerging communication will occur more readily when a child finds a display attractive and easy to understand, as compared to dull and challenging.

Given the research from the visual and cognitive sciences, it seems that the inclusion of human figures within VSDs holds promise as a powerful way to attract the visual attention and interest of viewers, including young children with complex communication needs. Support from within the field of AAC is offered by preliminary documentation of several cases in which infants and toddlers with complex communication needs (e.g., Down syndrome, cerebral palsy) demonstrated interest in VSDs that captured familiar animate figures, especially familiar faces engaged in social interaction (e.g., mom playing peek-a-boo; Light & Drager, 2008). These infants attended visually to these types of VSDs even at very young ages (e.g., prior to 12 months of age) and they used these VSDs to participate in social interactions. Light, et al. (in preparation) have proposed that these types of VSDs may provide a form of visual motherese for very young children. The visual display includes important animate elements (e.g., mom looking with wide eyes, surprised facial expression) that may engage the young child's interest and visual attention, thus maximizing the opportunity for communication and language learning.

Direct research is necessary to evaluate the role of humans in scenes across individuals of differing etiologies. While one might expect that children with strengths in social-pragmatic functioning (for instance, many children with Down syndrome) may be attracted to scenes containing people, what of individuals with social-pragmatic challenges? Initial studies using eyetracking technology have demonstrated that individuals with autism spectrum disorders show either unusual patterns of observation of human faces presented in isolation (Pelphrey, Sasson, Reznick, Paul, Goldman, & Piven, 2002) or reduced observation of human figures in scenes compared to matched participants without autism (Riby & Hancock, 2008, 2009). However, while examination of the data from Riby and Hancock's work (2008, 2009) does illustrate a reduction in the time individuals with ASD spent observing human figures, it is not wholly absent by any measure. Indeed, the individuals with autism did observe the human figures, albeit for briefer periods than their nondisabled peers. Thus, individuals with ASD were attending to the humans in the scenes, and it would be important to explore what kind of information was indeed being extracted. Similarly, Fletcher-Watson et al (2009) found that participants with autism spectrum disorders showed a preference to look at scenes containing people as compared to scenes that did not, in a free-viewing split-scene presentation.

Clearly, these issues are critical for clinical practice and the design of VSDs to support communication. Future research is needed to determine the effects of humans in VSDs on attention, comprehension, and communicative use by individuals with complex communication needs and to explore whether differences emerge across children with disabilities of different etiologies. Research is also needed to evaluate the impact across socio-cultural groups, as the effects could play out differently in different cultures. It may also be of interest to investigate cognitive visual processing patterns for animate stimuli other than humans (e.g., the effects of animals within scenes) and to investigate the effects of animacy across different VSD types (e.g., photographs, realistic line drawings, cartoons).

Depiction of an Event within a VSD

The humans in VSDs are not simply presented in isolation; rather they are engaged in meaningful events within the scenes presented (e.g., playing a tickle game, blowing bubbles, singing songs). The rationale underlying use of VSDs is that they capture the experiences children have had with the various concepts being represented. Thus, the oral language user learns the concept of dog, including the spoken word, while experiencing the events

surrounding the dog barking, or patting it, or being kissed by it (Nelson, 1986). The user of a VSD might similarly map the concept of “dog” through the VSD depiction of the event of patting and being kissed by the animal, with the relevant language concept embedded within the depicted event. On this logic, if VSDs take advantage of experience-based schemas, then the event itself seems a necessary compositional element to consider.

Depicting Events within VSDs

When photographic VSDs are designed to capture meaningful events within the lives of young children, the people within the scene are focused on the event that is occurring (e.g., the father and child may be looking at the ball and/or each other as they are playing catch). Because the photographs of these scenes (used as VSDs) capture a single moment in time, however, they do not contain as many of the rich attentional cues as would naturally occur in the actual social event. For instance, in the VSD in Figure 1, a single moment in the child's birthday experience is captured: The child is looking at the cake, and her reach with her arms is fairly general and nonspecific. In order to understand the VSD and use it to support communication, the user must attend to and recognize not only the child, but also the event (e.g., the presentation of the birthday cake). The fact that the user participated in the event herself might provide support (through episodic memory cued by the VSD) for her understanding of, and attention to, the key elements of the event, particularly the importance of the cake to the overall VSD.

In addition to depicting the experience/event itself, other elements of the VSD may support interpretation of more subtle aspects of the event. For instance, in Figure 2, the child is clearly referencing the contents of the bucket (a frog) by pointing. In actual interaction, this child would probably alternate gaze between the viewer/partner and the bucket, thus further cuing attention to the contents of the bucket. However, the photograph of the scene captures only one specific moment in the interaction so that the child is actually looking at the camera (and thus, the viewer), not the bucket. Missing is the dynamic sequence of events that in real life cues point of reference; e.g., pointing at the bucket and the additional powerful cue of the girl's eye gaze alternating between the bucket/frog and the partner. If children can infer the point of reference in such a photograph, they should have little difficulty with interpretation of the VSD and its specific programmable concepts such as “look at that!” or “what's that?” However, given the static nature of the photograph of the event, it is important to consider whether viewers of these static photographs understand the points of reference in the same way that they might if the events were actually occurring in a dynamic interaction.

Information about Related Issues in Visual and Cognitive Sciences

The ability to follow a partner's gaze or point during social interactions (receptive joint attention) has been studied extensively, and has been reported both to emerge relatively early in typical development and to be related to later language outcomes (e.g., Watt, Wetherby, & Shumway, 2006; Yoder & McDuffie, 2006). Over the last 15 years, there has been an increasing interest in how humans use visual cues like gaze, posture, and facial features to interpret the point of reference or subject of attention during real-time interactions (e.g., see Langton, Watt, & Bruce, 2000). There is surprisingly little research concerning viewers' ability to follow the point of reference of people in a static two-dimensional stimulus like a photograph, however. One exception is a study that examined whether or not nondisabled adult participants followed the gaze of a subject in a photograph (Fletcher-Watson et al., 2008). These researchers presented participants with a photograph containing a human subject who appeared to be looking at or scrutinizing something. Using eye-tracking technology, they investigated the participants' eye gaze and found that participants fixated on the item being scrutinized by the person in the photograph more than

would be expected by chance (i.e., proportional to the amount of gaze allocated to other objects around the subject). The participants clearly attended not only to the human subject, but also to the subject's point of reference (i.e., the object of attention).

How Might this Research Relate to VSDs and Individuals with Disabilities?

There remain a host of questions with respect to how to depict events in VSDs and the extent to which viewers should and can rely on static cues to determine point of reference in these scenes. It will be particularly critical to examine implications for VSDs across early childhood. At the earliest stages of development, infants may respond best to VSDs where the event is quite simple, with no demands to follow a point of reference. For example, in a VSD representing an early turntaking game of peekaboo, a mother could be photographed looking directly toward the camera (and hence toward the child using the VSD). The peek-a-boo interaction captured in the VSD is one of social closeness involving only the mother and the child; this VSD capitalizes on the presence of a person engaged in an event to attract the infant, but does not require the infant to determine an additional point of reference. As children develop joint attention skills, they may be better able to use VSDs that capture events involving shared activities such as those illustrated in the potential VSDs in Figures 1 and 2.

The research by Fletcher-Watson et al (2008) suggested that adults without disabilities are able to discern the point of reference in two-dimensional photographs of scenes. It is unclear whether these results will generalize to children and especially to children with complex communication needs, however. Young children may rely more heavily on a rich array of cues in order to discern point of reference than adults do. Given that static two-dimensional photographs of scenes may omit some of the cues available in dynamic natural interactions, children may have more difficulty determining the point of reference in VSDs than adults do, especially in the early stages of development. Children with developmental disabilities might be even more vulnerable to reductions in the cues for point of reference and may have more difficulty following the subtle cues captured in static form in photographs. Moreover, not all children develop joint attention as expected. Infants who are later diagnosed with autism often demonstrate difficulties in receptive joint attention; moreover their development of joint attention (or lack thereof) is a strong predictor of later language and communication outcomes (e.g., Charman, Baron-Cohen, Swettenham, Baird, Drew, & Cox, 2003; Sullivan, Finelli, Marvin, Garrett-Mayer, Bauman, & Landa, 2007; Toth, Munson, Meltzoff, & Dawson, 2006; Warreyn, Roeyers, Van Wetswinkel, & De Groote, 2007; Wetherby, 2006). Indeed, a follow-up study by Fletcher-Watson et al. (2009) examined whether older individuals with ASDs followed the gaze in the photographs just described (i.e., photographs containing a human subject who appeared to be looking at or scrutinizing something). They found that while the participants with ASD did actually look at the human figure in the photo, they did not attend to the item being scrutinized any longer than would be expected by chance, in contrast to the nondisabled participants. A better understanding of visual attention and processing within and across populations of individuals who use AAC is therefore urgently needed.

The potential vulnerabilities in developing joint attention make it important to consider approaches to designing VSDs that support children in determining the point of reference. Is it possible to manipulate the 2-dimensional format (e.g., photograph) to emphasize the referential point more clearly? What is the impact of using video to highlight a greater range of dynamic cues? We noted in earlier sections that (a) VSDs enable use of highly familiar and personalized images, and (b) the literature in visual cognitive science/neuroscience suggests that the presence of familiar context may in part underlie the very rapid processing of naturalistic scenes. Might clinicians promote understanding of point-of-reference through images that include such referencing by familiar partners or, in fact, the child him/herself?

These are questions that must be answered if clinicians are to design visual scene displays that capitalize on the visual-cognitive processing advantages inherent in naturalistic scenes to support children's participation in social interactions. Future research is required to investigate these issues to determine how young children with complex communication needs develop point of reference skills over time and how they utilize these skills to make sense of VSDs for communication purposes.

Conclusions

Research on visual cognitive processing of photographs of naturalistic scenes demonstrates that humans are able to process naturalistic scenes and recognize and classify the overall context and constituent elements (i.e., people and objects) very rapidly. Such processing of scenes may be facilitated by the fact that scenes preserve the conceptual and visual relationships between elements as experienced in real life. Processing may be bolstered by real world experiences of the events, that allow for the prediction of expected elements within the scene. Finally, certain elements appear to hold unique powers of attraction in the processing of natural scenes, most particularly human figures in these scenes.

VSDs may have the potential to engage the same processes that have been reported to facilitate the processing of naturalistic scenes in the visual cognitive science literature. If so, that body of research could have potential implications for AAC, especially interventions with young children with complex communication needs. Since these children are at the early stages of development, they may be especially vulnerable when AAC systems impose too great a processing load. The research in cognitive visual processing suggests that the processing load imposed by AAC systems may be reduced through VSDs, especially those that capture meaningful social events involving humans.

Limitations and Future Research Directions

Readers are cautioned that the considerations proposed in this paper, while generally very well-researched in visual cognitive neuroscience, have not yet been examined within the field of AAC. Very little work has been done developmentally (with younger children) or from the perspective of individuals with disabilities. The current research therefore does not yet directly reflect the population of clients who might use AAC. Nonetheless, an understanding of general principles of human visual processing may offer a starting point for the field of AAC. Future research is obviously required to investigate whether these principles are influenced by developmental status (age) or the presence of disability (cf. Higginbotham, 1995). Indeed, one outcome we hope to promote with this analysis is the initiation of future research by researchers and clinicians within AAC in conjunction with researchers in the visual cognitive sciences.

We have discussed only a small number of the possible considerations that may be important, within the context of the visual-perceptual demands or supports that VSDs present. These considerations must always be placed within the context of functional communication that is the goal of aided AAC intervention. It is of paramount importance to ensure that the content depicted within VSDs is motivating, meaningful, and relevant to the children with complex communication needs who use them. Even a display that is perfect in visual composition will be limited in functionality if its content is irrelevant or unmotivating. Our argument is simply that motivating content will be more effective if placed within a display that is well-matched to the user's visual and cognitive processing.

Clearly, there may be other visual structural dimensions that may require investigation as well. It is critical to examine the extent to which research from basic studies in visual cognitive science and related fields applies when extended to VSDs intended to be used for

purposes of communication. Additionally, it will be critical to extend the evidence base from nondisabled adults to younger children, particularly those with complex communication needs, including those with concomitant visual impairments. Whether recommendations will vary depending on etiology is another key point that warrants examination. Future research is also required to begin to untangle the effects of specific elements within visual scenes and the effects of these elements on processing. Finally, a thorough consideration of issues related to iconicity is warranted, in particular, examination of the features that contribute to guessability – such as the type of symbol, the abstractness of the concept being represented, the contextual cues that support interpretation, and the familiarity of the concept to the user.

This research agenda will require the expertise of multiple disciplines; the extensive knowledge of human processing from other fields offers a strong basis from which to begin examination of these types of issues within the area of AAC. With AAC systems that are designed to match the needs and skills of young children, professionals will be better able to improve outcomes for young children with complex communication needs.

Acknowledgments

This research was supported in part through three grants: (1) the Communication Enhancement Rehabilitation Engineering Research Center (AAC_RERC), a virtual research center that is funded by the National Institute on Disability and Rehabilitation Research (NIDRR) under grant H133E030018, and (2) grant #P01 HD25995 from the National Institute of Child Health and Human Development (NICHD). The opinions contained in this publication are those of the grantees and do not necessarily reflect those of the granting agencies.

References

- Barton A, Sevcik RA, Ronski MA. Exploring visual-graphic symbol acquisition by preschool-age children with developmental and language delays. *Augmentative and Alternative Communication*. 2006; 22:10–20. [PubMed: 17114155]
- Biederman, I. On the semantics of a glance at a scene.. In: Kubovy, M.; Pomerantz, JR., editors. *Perceptual Organization*. Erlbaum; Hillsdale, NJ: 1981. p. 213-253.
- Blackstone, S. What are Visual Scene Displays?. *Augmentative Communication News*, 1. 2005. http://www.imakenews.com/aac-lerc/e_article000344804.cfm?x=b11,0,w
- Bloomburg K, Karlan G, Lloyd L. The comparative translucency of initial lexical items represented by five graphic symbols and sets. *Journal of Speech and Hearing Research*. 1990; 33:717–725. [PubMed: 2148789]
- Bower, TGR. *A Primer of Infant Development*. Freeman; San Francisco: 1977.
- Bondy A, Frost L. The picture exchange communication system. *Behavior Modification*. 2001; 25:725–745. [PubMed: 11573337]
- Bopp KD, Brown KE, Mirenda P. Speech-language pathologists' roles in the delivery of positive behavior support for individuals with developmental disabilities. *American Journal of Speech-Language Pathology*. 2004; 13:5–19. [PubMed: 15101810]
- Bracco F, Chiorri C. People have the power: Priority of socially relevant stimuli in a change detection task. *Cognitive Processes*. 2009; 10:41–49.
- Bruner, J. *Child's talk: Learning to use language*. Norton; New York: 1983.
- Braun J. Natural scenes upset the visual appletart. *Trends in Cognitive Sciences*. 2003; 7:7–9. [PubMed: 12517351]
- Charman T, Baron-Cohen S, Swettenham J, Baird G, Drew A, Cox A. Predicting outcomes in infants with autism and pervasive developmental disorder. *International Journal of Language and Communication Disorders*. 2003; 38:265–285. [PubMed: 12851079]
- Crouzet SM, Kirchner H, Thorpe SJ. Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision*. 2010; 10:1–17. [PubMed: 20465335]
- Culp DM, Ambrosi DM, Berniger TM, Mitchell JO. Augmentative communication aid use – a followup study. *Augmentative and Alternative Communication*. 1986; 2:19–24.

- Dietz, A.; McKelvey, M.; Hux, K.; Beukelman, D.; Wallace, S.; Weissling, K. Integrating contextual relevant visual scenes into aphasia interventions.. Paper presented at the American Speech Language Hearing Association Convention; Miami FL. 2006, November;
- Drager K, Light J, Curran-Speltz J, Fallon K, Jeffries L. The performance of typically developing 2 ½-year-olds on dynamic display AAC technologies with different system layouts and language organizations. *Journal of Speech, Language and Hearing Research*. 2003; 46:298–312.
- Drager K, Light J, Carlson R, D'Silva K, Larsson B, Pitkin L, Stopper G. Learning of dynamic display AAC technologies by typically developing 3-year olds. *Journal of Speech, Language and Hearing Research*. 2004; 47:1133–1148.
- Fei-Fei L, VanRullen R, Koch C, Perona P. Why does natural scene categorization require little attention? Exploring attentional requirements for natural and synthetic stimuli. *Visual Cognition*. 2005; 12:893–924.
- Fletcher-Watson S, Findlay JM, Leekam SR, Benson V. Rapid detection of person information in a naturalistic scene. *Perception*. 2008;571–583. [PubMed: 18546664]
- Fletcher-Watson S, Leekam SR, Benson V, Frank MC, Findlay JM. Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*. 2009; 47:248–257. [PubMed: 18706434]
- Fuller D. Initial study into the effects of translucency and complexity on the learning of Blissymbols by children and adults with normal cognitive abilities. *Augmentative and Alternative Communication*. 1997; 13:30–39.
- Fuller D, Lloyd L. A study of physical and semantic characteristics of a graphic symbol system as predictors of perceived complexity. *Augmentative and Alternative Communication*. 1987; 3:26–35.
- Gliga T, Elsabbagh M, Andravizou A, Johnson M. Faces attract infants' attention in complex displays. *Infancy*. 2009; 14:550–562.
- Harris MD, Reichle J. The impact of aided language stimulation on symbol comprehension and production in children with moderate cognitive disabilities. *American Journal of Speech Language Pathology*. 2004; 13:155–167. [PubMed: 15198634]
- Higginbotham DJ. Use of nondisabled subjects in AAC research: Confessions of a research infidel. *Augmentative and Alternative Communication*. 1995; 11:2–5.
- Hopkins KA. Why do babies find faces attractive? *Australian Journal Childhood*. 1980; 5:25–28.
- Hurlbut BI, Iwata BA, Green JD. Nonvocal language acquisition in adolescents with severe physical disabilities: Blissymbol versus iconic stimulus formats. *Journal of Applied Behavior Analysis*. 1982; 15:241–248. [PubMed: 6181049]
- Jagaroo V, Wilkinson KM. Further considerations of visual cognitive neuroscience for aided AAC: The potential role of motion perception systems in maximizing design display. *Augmentative and Alternative Communication*. 2008; 24:29–42. [PubMed: 18938756]
- Johnson JM, Inglebrete E, Jones C, Ray J. Perspectives of speech-language pathologists regarding success versus abandonment of AAC. *Augmentative and Alternative Communication*. 2006; 22:85–99. [PubMed: 17114167]
- Johnston SS, Reichle J. Designing and implementing interventions to decrease problem behaviors. *Language, Speech, and Hearing Services in Schools*. 1993; 24:225–235.
- Joseph RM, Brandon K, Connolly C, Wolfe J, Horowitz TS. Why is visual search superior in autism spectrum disorder? *Developmental Science*. 2009; 12:1083–1096. [PubMed: 19840062]
- Kayser C, Kording KP, König P. Processing of complex stimuli and natural scenes in the visual cortex. *Current Opinion in Neurobiology*. 2004; 14:468–473. [PubMed: 15302353]
- Klin A, Jones W, Schultz R, Volkmar F, Cohen D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*. 2002; 59:809–816. [PubMed: 12215080]
- Langton S, Watt RJ, Bruce V. Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*. 2000; 4:50–59. [PubMed: 10652522]
- Light, J.; Drager, K. Evidence-based AAC interventions to build language and communication skills with infants, toddlers, and preschoolers.. Seminar presented at the ISAAC conference; Montreal, Canada. 2008, August;

- Light J, Drager K, McCarthy J, Mellott S, Parrish C, Parsons A, Rhoads S, Ward M, Welliver M. Performance of typically developing four and five year old children with AAC systems using different language organization techniques. *Augmentative and Alternative Communication*. 2004; 20:63–88.
- Light J, Lindsay P. Cognitive science and augmentative and alternative communication. *Augmentative and Alternative Communication*. 1991; 7:186–203.
- Liss M, Saulnier C, Fein D, Kinsbourne M. Sensory and attention abnormalities in autistic spectrum disorders. *Autism: The International Journal of Research & Practice*. 2006; 10:155–172. [PubMed: 16613865]
- Mack SC, Eckstein MP. Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of Vision*. 2011; 9:1–16.
- Mann T, Walker P. Autism and a deficit in broadening the spread of visual attention. *Journal of Child Psychology & Psychiatry & Allied Disciplines*. 2003; 44:274–284.
- McKelvey ML, Hux K, Dietz A, Beukelman DR. Impact of personal relevance and contextualization on word-picture matching by people with aphasia. *American Journal of Speech-Language Pathology*. 2010; 19:22–33. [PubMed: 20139353]
- Mirenda, P.; Iacono, T., editors. *Autism spectrum disorders and AAC*. Paul H. Brookes; Baltimore: 2009.
- Mirenda P, Locke P. A comparison of symbol transparency in nonspeaking persons with intellectual disabilities. *Journal of Speech and Hearing Disorders*. 1989; 54:131–140. [PubMed: 2709831]
- Mizuko M, Reichle J. Transparency and recall of symbols among intellectually handicapped adults. *Journal of Speech and Hearing Disorders*. 1989; 54:627–633. [PubMed: 2478760]
- Mizuko M, Reichle J, Ratcliff A, Esser J. Effects of selection techniques and array sizes on short-term visual memory. *Augmentative and Alternative Communication*. 1994; 10:237–244.
- Nelson, K. *Event knowledge: Structure and function in development*. L. Erlbaum Associates; Hillsdale, N.J.: 1986.
- New J, Cosmides L, Tooby J. Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences*. 2007; 104:16598–16603.
- Oliva A, Torralba A. The role of context in object recognition. *Trends in Cognitive Sciences*. 2007; 11:520–527. [PubMed: 18024143]
- Peelen M, Fei-Fei L, Kastner S. Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature (letters)*. 2009; 460:94–97.
- Pelphrey KA, Sasson NJ, Reznick JS, Paul G, Goldman BD, Piven J. Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*. 2002; 32:49–261.
- Phillips B, Zhao H. Predictors of assistive technology abandonment. *Assistive Technology*. 1993; 5:36–45. [PubMed: 10171664]
- Plaisted K, O'Riordan M, Baron-Cohen S. Enhanced discrimination of novel, highly similar stimuli by adults with autism during a perceptual learning task. *Journal of Child Psychology and Psychiatry*. 1998; 39:765–775. [PubMed: 9690939]
- Plaisted K, Swettenham J, Rees L. Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of Child Psychology and Psychiatry*. 1999; 40:733–742. [PubMed: 10433407]
- Rensink RA, O'Regan J, Clark J. To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*. 1997; 8:368–373.
- Riby D, Hancock PJB. Viewing it differently: Social scene perception in Williams syndrome and Autism. *Neuropsychologia*. 2008; 46:2855–2860. [PubMed: 18561959]
- Riby D, Hancock PJB. Looking at movies and cartoons: Eye tracking evidence from Williams Syndrome and autism. *Journal of Intellectual Disability Research*. 2009; 53:169–181. [PubMed: 19192099]
- Romski, MA.; Sevcik, RA. *Breaking the speech barrier: Language development through augmented means*. Brookes; Baltimore: 1996.
- Smilek D, Birmingham E, Cameron D, Bischof W, Kingstone A. Cognitive ethology and exploring attention in real-world scenes. *Brain Research*. 2006; 1080:101–119. [PubMed: 16480691]

- Sullivan M, Finelli J, Marvin A, Garrett-Mayer E, Bauman M, Landa R. Response to joint attention in toddlers at risk for autism spectrum disorders: a prospective study. *Journal of Autism and Developmental Disorders*. 2007; 37:37–48. [PubMed: 17216332]
- Toth K, Munson J, Meltzoff AN, Dawson G. Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation, and toy play. *Journal of Autism & Development Disorders*. 2006; 36:993–1005.
- Venturino M, Gagnon DA. Information tradeoffs in complex stimulus structure: Local and global levels in naturalistic scenes. *Perception and Psychophysics*. 1992; 52:425–436. [PubMed: 1437475]
- Warreyn P, Roeyers H, Van Wetswinkel U, De Groote I. Temporal coordination of joint attention behavior in preschoolers with autism spectrum disorder. *Journal of Autism and Developmental Disorders*. 2007; 37:501–512. [PubMed: 16915534]
- Watt N, Wetherby A, Shumway S. Prelinguistic predictors of language outcome at 3 years of age. *Journal of Speech, Language, and Hearing Research*. 2006; 49:1224–1237.
- Wetherby, AM. Understanding and measuring social communication in children with autism spectrum disorders.. In: Charman, T.; Stone, W., editors. *Social and communication development in autism spectrum disorders: Early identification, diagnosis, and intervention*. Guilford Press; New York: 2006. p. 3-34.
- Wetherby, A.; Prizant, B. *The Communication and Symbolic Behavior Scales*. Paul H. Brookes; Baltimore: 1993.
- Wilhelm H, Lovaas O. Stimulus overselectivity: a common feature in autism and mental retardation. *American Journal of Mental Deficiency*. 1976; 81:26–31. [PubMed: 941930]
- Wilkinson KM, Hennig S. State of the art and current recommended practice in augmentative and alternative communication. *Mental Retardation and Developmental Disabilities Research Reviews*. 2007; 13:58–69. [PubMed: 17326111]
- Wilkinson KM, Jagaroo V. Contributions of visual cognitive neuroscience to AAC display design. *Augmentative and Alternative Communication*. 2004; 20:123–136.
- Wilkinson KM, Light J. Observation of humans in naturalistic visual scenes: Implications for the design of symbols and displays for use in aided AAC interventions. *Journal of Speech-Language-Hearing Research (Language)*. 2011; 54:1644–1657.
- Yoder, P.; McDuffie, AS. Treatment of responding to and initiating joint attention.. In: Charman, T.; Stone, W., editors. *Social and communication development in autism spectrum disorders*. Guilford Press; NY: 2006. p. 117-142.

Figure 1.
Example of a visual scene display for a young child's birthday party.

Figure 2.
A visual scene display of a child referencing something within the VSD (point of reference)