AJSLP

Research Article

What's in a Photograph? The Perspectives of Composition Experts on Factors Impacting Visual Scene Display Complexity for Augmentative and Alternative Communication and Strategies for Improving Visual Communication

Kevin M. Pitt[a] and John W. McCarthy[b]

Purpose: Visual scene displays (VSDs) can support augmentative and alternative communication (AAC) success for children and adults with complex communication needs. Static VSDs incorporate contextual photographs that include meaningful events, places, and people. Although the processing of VSDs has been studied, their power as a medium to effectively convey meaning may benefit from the perspective of individuals who regularly engage in visual storytelling. The aim of this study was to evaluate the perspectives of individuals with expertise in photographic and/or artistic composition regarding factors contributing to VSD complexity and how to limit the time and effort required to apply principles of photographic composition.
Method: Semistructured interviews were completed with 13 participants with expertise in photographic and/or artistic composition.
Results: Four main themes were noted, including (a) factors increasing photographic image complexity and decreasing cohesion, (b) how complexity impacts the viewer, (c) composition strategies to decrease photographic image complexity and increase cohesion, and (d) strategies to support the quick application of composition strategies in a just-in-time setting. Findings both support and extend existing research regarding best practice for VSD design.
Conclusions: Findings provide an initial framework for understanding photographic image complexity and how it differs from drawn AAC symbols. Furthermore, findings outline a toolbox of composition principles that may help limit VSD complexity, along with providing recommendations for AAC development to support the quick application of compositional principles to limit burdens associated with capturing photographic images.
Supplemental Material: https://doi.org/10.23641/asha.15032700

After taking a digital photo, it is easy to review and decide whether to keep or retake it. The decision to keep or discard an image can be made for a variety of reasons. In some cases, the subject of the photo may have been moving or is not in the frame. In other cases, a more subjective decision may be made about whether or not it is a "good" photo, and explanations of why someone did or did not like a photo may be harder to obtain. Having a good photo can inspire, quickly engender feelings or, ideally, capture the moment. Photos are used frequently in social media and for quick communication of ideas online, but despite the seemingly disposable presence of photos, decision making about their suitability for communicative purposes need not be a solely intuitive process.

[a]Department of Special Education and Communication Disorders, University of Nebraska–Lincoln
[b]Division of Communication Sciences and Disorders, Ohio University, Athens
Correspondence to Kevin M. Pitt: kevin.pitt@unl.edu

Editor-in-Chief: Julie Barkmeier-Kraemer
Editor: Katherine C. Hustad
Received November 13, 2020
Revision received February 12, 2021
Accepted April 21, 2021
https://doi.org/10.1044/2021_AJSLP-20-00350

Decision Making for Images in Education and Health Care

In the medical field, there is research exploring what makes a good photo for documentation and educational purposes. Specifically, research in the areas of ophthalmology (Mukherjee & Nair, 2012), orthopedics (Uzun et al.,

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

2014), and dermatology (Muraco, 2020) have identified a range of positive compositional factors for photography, including filling the frame with the object of interest while providing enough surrounding information to identify anatomical relationships, providing sufficient lighting and exposure, keeping objects of interest in focus, limiting focal object cutoff, and avoiding unnatural angles and perspectives. When more than one image is presented simultaneously, there are additional considerations. For example, in examining how eye tracking could be used to more precisely assess comprehension in individuals with aphasia, Heuer and Hallowell (2007) found that physical stimulus properties such as color, orientation, size, and luminance may impact attentional patterns of those with aphasia during viewing of multiple-choice arrays. Therefore, pictorial arrays needed to be balanced across all features with careful design principles (Heuer et al., 2017).

In working with people with aphasia, Beukelman et al. (2015) suggested that highly contextualized photographs were best for use in restorative and compensatory approaches. These high-context or contextually rich photos had clear relationships provided between personally relevant people and objects in a meaningful environment. Therefore, carefully chosen photos could provide a scaffold for therapy activities or the basis for a co-constructed conversation (Hux et al., 2010). The communicative benefits of contextually rich photos are thus the basis for visual scene displays (VSDs) in augmentative and alternative communication (AAC).

VSDs

Using VSDs can support communication success for adults and children with complex communication needs. These photographic images provide information about the natural environment in which people, places, and objects occur (e.g., during a child's sporting event) and embed hotspots within the scene store an associated communicative message (e.g., Beukelman & Light, 2020). The provision of context helps clarify image content, supporting language abilities (Beukelman et al., 2015), and takes advantage of individuals' natural ability to quickly process scenes, which possibly lowers cognitive burdens associated with interpretation of the visual display (Brown et al., 2019).

In validating the processing demands of VSDs, research has focused on a range of factors, such as the number of items within the scene (Wilkinson et al., 2012), depiction of shared events and task engagement (i.e., individuals within the scene are engaging in the depicted event; Thiessen et al., 2017; Wilkinson et al., 2012), background elements during a shared activity (O'Neill et al., 2019), number of directional lines (Wilkinson et al., 2012), the inclusion of human figures (Thiessen et al., 2014; Wilkinson et al., 2012; Wilkinson & Light, 2014), and familiarity and personalization (Dietz et al., 2014; McKelvey et al., 2010). The research generally focuses on attention demands with different variations in a scene. For example, in a preliminary study, Wilkinson and Light (2011) tracked the visual attention patterns of 19 college students aged 18–22 years when looking at eight different visual scene images. Images included people in conditions where the people were relatively small or where other objects, because of their size, would appear to compete for attention. Their findings indicated that participants had a fast and enduring allocation of their visual attention to people regardless of their relative size or competing imagery. Therefore, the evidence suggests that scene elements and design may impact the gaze patterns of individuals viewing VSDs.

AAC research has also considered the processing demands of grid-based displays. A number of factors that impact outcomes for grid-based AAC displays are identified, such as clustering symbols by internal color (e.g., Wilkinson et al., 2008), use of background (Thistle & Wilkinson, 2017) and foreground color (Thistle & Wilkinson, 2009), along with the influence of matrix size (Thistle, 2019), symmetry and symbol orientation (Wilkinson & Jagaroo, 2004), motion (Jagaroo & Wilkinson, 2008), the incorporation of text alongside symbols (Brown et al., 2015), space and symbol arrangement such as spatial cuing (Light, Wilkinson, et al., 2019), organizing by word class category (Thistle & Wilkinson, 2017), organizing by emotional category (Wilkinson & Snell, 2011), and considering left–right hemisphere processing bias (Wilkinson & Jagaroo, 2004). In addition, studies evaluating how children draw symbol concepts (e.g., eat) identified that children are more likely to depict concepts using whole shapes (Light & Drager, 2007; McCarthy et al., 2018), which, when in context, may help support accurate symbol identification (Worah et al., 2015). These important findings regarding how the principles of the visual cognitive neurosciences can inform display design have led to initial guidelines for clinical practice (e.g., Light, Wilkinson, et al., 2019), which may lower cognitive–perceptual loads and task difficulty.

Similar to grid displays, conceptualizing the notion of photographic image complexity and how to compose VSD can help support improved AAC performance (Light, Wilkinson, et al., 2019). However, while there may be some overlap, what makes a naturalistic VSD image complex or difficult to understand may differ from what makes a symbol complex in a grid system (Wilkinson et al., 2012). Therefore, as digital photographs become easier than ever to obtain for communication purposes (Brown & Thiessen, 2018), there is still room for consideration of what makes a "good" VSD photograph and how complexity and compositional factors can impact VSD-based communication.

Maximizing the Communicative Power of a Scene to Create Meaning

In the context of scene viewing, both saliency and meaning factors impact how one reads an image and interprets meaning. Saliency refers to low-level (bottom-up) image features, such as luminance, contrast, and color, which allow objects to stand out from their surroundings. These salient focal points may attract our attention, independent of top-down processes related to overall scene meaning (Itti & Koch, 2001; Wu et al., 2014). In more detail, scene regions that are uniform along salient features are identified as uninformative, with scene regions that differ along salient features identified as informative (Henderson & Hayes, 2018). For instance, scene viewers may be attracted to the brightest element(s) in the scene, as the brightness difference causes a change in salience. In this manner, saliency helps provide a framework for how our attention is guided to scene elements (Itti & Koch, 2000; Itti et al., 1998; Liu & Gleicher, 2006). In contrast, how scenes are viewed is also impacted by a top-down process, which guides our attention to areas of the scene that are most semantically informative and relevant, based on our world knowledge, general scene schemas (or template), and the scene itself (Henderson & Hayes, 2018).

Previous research discusses important scene elements and has established how VSDs can reduce processing challenges versus traditional grid methods. However, a synthesis is lacking that informs how to maximize the effectiveness of a photograph to capture a moment. In line with prior research utilizing principles of the visual cognitive neurosciences, the field of AAC can benefit from incorporating the perspectives of outside disciplines into conversations related to AAC innovation and service delivery. Working with visual scenes is something professionals with a background in photography and/or artistic principles of composition have been doing for years. Principles of photographic and art composition with an orientation to convey a specific message, story, or perspective may support findings from the visual cognitive sciences by helping understand what makes scenes more complex and how to limit scene complexity by increasing image structure, decreasing distraction, and drawing the viewers' attention to key scene elements (Peterson, 2003). In addition, photographic and composition professionals may shed new light on how compositional principles can be easily and quickly applied to support rapid image acquisition and support just-in-time programming of VSDs during daily interactions (Light, McNaughton, & Caron, 2019). Therefore, the aim of this study was to identify the perspectives of compositional experts who possess a range of experiences and expertise to help guide how VSD can be designed to decrease cognitive load and distraction while increasing individuals' efficiency during AAC use.

Method

Participants

Approval from the institutional review board at the University of Nebraska–Lincoln and Ohio University was provided prior to study commencement. Thirteen participants (n = 7 women, M = 42.5 years, SD = 9.8, range: 23–55) completed the study. Participants had an average of 17.5 years (SD = 10.7, range: 4–38) of experience with photographic and/or artistic composition. A range of backgrounds related to composition were targeted for inclusion (see Table 1 for participant demographics). Furthermore, one participant had a brother diagnosed with Down syndrome and had spent time teaching individuals with this diagnosis how to use a camera to express themselves. Prior to the interviews being conducted, a short presentation about VSDs was delivered to participants to ensure all had a common frame of reference. The presentation included a general description about the field of AAC, VSDs, and hotspots. Furthermore, a link to a public video showing VSD use was provided for reference. The video was posted on an AAC-focused university's webpage and included scenes of a trip that both included and excluded people. Those experts incorporated into this investigation had a minimum of 2 years of experience related to composition, felt comfortable discussing compositional principles, and self-determined their experience could be extended to inform VSD design.

Materials

The interview guide (see Supplemental Material S1) was developed to fill gaps in current literature regarding what contributes to VSD complexity and applications of photographic and artistic composition, discussing (a) what makes a photographic image complex; (b) possible impacts of image complexity of the viewer; (c) how the principles of photographic and/or artistic composition can help limit VSD complexity, decrease distraction, and highlight key scene elements; and (d) what camera settings could be used to emphasize these compositional elements automatically to decrease stakeholder burdens with image collection.

One additional interview question discussing physical access techniques for VSDs was removed from the provided interview protocol, and data will be presented elsewhere. During interviews, a focus was placed on compositional principles that may be quickly applied to support just-in-time programming of naturally occurring communication events.

Procedure

Interview and data analysis procedures were based upon detailed reports of qualitative methods currently used in the field of AAC (e.g., Hajjar et al., 2016; O'Neill & Wilkinson, 2020). All interviews were completed in a single session via video conferencing (n = 12) or telephone (n = 1). Interviews lasted approximately 45 min and followed an interview guide to help ensure systematic data collection. Furthermore, at the beginning of each interview, the lead author reviewed the VSD presentation with participants and answered any related questions. Only once all participants' questions were answered and they felt comfortable did interviews commence. During interviews, some participants showed personal or web-based images to support communication about composition principles. The interviewer asked follow-up questions and requested clarifications to help ensure accurate coding and explore participant insights in more detail. In addition, notes were taken by the interviewer during data collection to prompt follow-up questions and support later analysis by recording key quotes and phrases, along with initial interview themes. The lead author completed all interviews between January 2020 and May 2020.

Table 1. Participant demographics.

Participant	Age (years)	Gender	Years of experience	Primary areas of expertise
P1	38	F	21	Photography
P2	48	M	10	Light use in composition
P3	49	F	23	Photography/film making
P4	31	M	8	Photography, drawing, printmaking
P5	50	F	32	Photography
P6	42	M	14	Human attention during scene perception
P7	30	F	15	Photography
P8	55	M	38	Visual art
P9	23	F	5	Journalism/photography
P10	43	M	13	Cinematography
P11	50	F	4	Commercial photography
P12	53	M	30	Visual photography
P13	40	F	14	Film

Note. F = female; M =male.

Data collection and analysis were ceased at saturation, when collecting new data no longer revealed new insights, and all members of the research team (including two speech-language pathologists, one with a degree in photography and video and one with experience in photography, plus one graduate student in speech-language pathology with a background in visual processing) agreed topics had been discussed in sufficient detail (Saunders et al., 2018). Specifically, data collection finished when no new codes were identified for three consecutive participants and the research team reached 100% consensus that the themes were fully comprehensible.

Data Analysis

All of the interviews were audio-recorded and transcribed verbatim, including language features such as laughing and pausing, by a trained graduate assistant who also ensured the interview script had been followed. After initial transcription, a second graduate assistant checked transcription accuracy, with discrepancies discussed to 100% consensus. Following transcription, files were imported into NVivo software, which allows for organization and analysis of qualitative data, such as interview transcripts (QSR International, 2018). Using a grounded theory approach (Gibbs, 2008), interview themes were grouped by means of NVivo's coding features using a constant comparison approach through which new data were incorporated into the existing coding structure, with new codes added as new information emerged (Creswell, 2012). Following coding, a codebook detailing the four major themes, 20 subthemes, and 52 codes identified was developed (see Supplemental Material S2). Following codebook generation, the lead author and trained graduate assistant reevaluated all interview transcripts to provide a subjective assessment of codebook consistency. Discrepancies in the codebook were then discussed among all team members until a final consensus was reached. Following consensus, an evaluation of reliability was conducted. In more detail, 23% (three of 13) of the interview transcripts were coded separately by a trained research assistant using the codebook. Based upon the procedures of O'Neill and Wilkinson (2020), the large number of codes identified in this study decreases the likelihood of coding agreements occurring by chance. Therefore, percent agreement (Syed & Nelson, 2015) was chosen for evaluating reliability for the randomly identified transcripts. Intercoder reliability was performed independently by a trained graduate assistant until more than 80% accuracy was achieved at the level of the code (McHugh, 2012; O'Neill & Wilkinson, 2020). For our investigation, an average of 94% (range: 87%–100%) intercoder reliability was achieved across the selected transcripts.

Credibility Indicators

Throughout the study, multiple techniques were used to further ensure data quality and credibility, including member checking, peer debriefing and review, and triangulation (Creswell, 2012; Gibbs, 2008).

Member Checking

Member-checking procedures were completed both during and following the interview. During the interview, member checking was completed by the interviewer by (a) requesting participants to elaborate on unclear statements during the interview process and (b) providing summary statements to help ensure correct understanding during the interview. Following each interview, a discussion summary was sent to each interviewee so they could confirm that their ideas were represented in our evaluation. Twelve of the 13 participants responded to the request, all indicating agreement with summary content.

Peer Debriefing, Review, and Triangulation

The second author provided peer review of study methods across the length of the investigation. In addition, to confirm study findings were consistent with current theories regarding visual composition, following data analysis, a peer briefing procedure was completed (Brantlinger et al., 2005). The peer debriefer, who is currently completing their PhD in media communications, had 15 years of experience related to composition. The peer debriefer agreed that study findings were consistent with current theories regarding visual composition. Finally, a triangulation methodology was incorporated by utilizing a team approach (including both authors and a trained research assistant) during data analysis to help decrease the possibility of lead author bias.

Procedural Reliability

Procedural reliability was tracked via spreadsheet for study procedures related to transcription, distribution of the VSD presentation, and distribution of participant member checking documents. All procedures were completed with 98% accuracy as one participant (P11) indicated they did not receive the VSD presentation prior to beginning the interview. In this case, the interview only continued after the presentation was discussed, all their questions were answered, and they indicated they still wished to continue with the interview.

Results

An overview of identified themes and subthemes is provided in Table 2, with the full codebook provided in Supplemental Material S2. The following results are organized by theme and subtheme.

Factors Increasing Image Complexity and Decreasing Cohesion

Saliency Principles Are Not Used Purposefully

Twelve participants outlined how image complexity may be increased when similar (analogous) colors and shades are close together or overlap, or color is utilized in a non-purposeful manner. For example, P2 stated:

> Let's say that we've got a photograph that shows a bunch of kale next to some romaine lettuce next to maybe some broccoli, we've got all these green objects. There's not going to be anything that stands out because green does not help us isolate what we're looking at, even different shades of green really don't help us. Now, we have to start looking at line and start to identify objects by their shape and that takes more time.

Regarding similar colors, another participant described that an individual's color vision should be taken into account, as those with red and green color blindness may perceive these colors as similar. Seven participants noted that images in which everything is equal in brightness or which include a large amount of shadow may be more distracting, decreasing image unity and making it difficult to read, as one participant summarized:

> If something that doesn't have enough light it can be under exposed and suddenly you've lost information in the shadows. Um, we tend to want to find that information and we want to see that information. A good example is in your horror movies, a lot of times your monster is lurking in the dark and you can't necessarily identify him.

Leading lines not emphasizing the focal object(s) were discussed by four participants to increase complexity. For instance, P9 stated, "If there's a cord running through the image, then my eyes are, whether I'm conscious of it or not, being led out." Other areas participants identified as

Table 2. Themes and subthemes.

Theme	Subtheme
Factors increasing image complexity and decreasing cohesion	Saliency principles are not used purposefully. A unifying theme, context, or consistency with one's image schema is incomplete or missing.
	The number and completeness of elements
	Age, cognition, environment, and interests can potentially compound complexity.
How complexity may impact the viewer	Complexity decreases accessibility and increases difficulty in understanding image content.
	Complexity may be mitigated when an individual's interests are activated.
Composition strategies to decrease photographic image complexity and increase cohesion	Composition is not a one-size-fits-all approach; composition principles provide a toolbox to support photographic communication.
	Use of contrast (light and dark)
	Use of contrast (color)
	Structural principles
	Shape and space
	Leading line toward focal object or into frame
	Scale
	Focus
	Facial features should be in focus and well lit.
	Not distorting reality to support consistency with the viewer's image schema and support interpretation of meaning
Strategies to support the quick application of composition strategies in a just in time setting	Utilizing automatic grid lines and frames
	Automatic camera settings, especially aperture priority
	Touch screen apps for focus, light, and depth of field
	4 and 4.6k cameras and editing software

possibly increasing image complexity included (a) including multiple items as a single depth, like a police lineup where all suspects are arranged in a manner of equal importance, or having multiple items all at various depths, noted by three participants; (b) a lack of image structure, noted by four participants; and (c) equal or minimal use of focus, as noted by three participants, with P1 describing "one other thing that's kind of particular to photography also is the idea of focus, so if everything is equally sharp, everything is equally important."

A Unifying Theme, Context, or Consistency With One's Image Schema Is Incomplete or Missing

A total of 10 individuals discussed how factors that may contradict an individual's image schema, distort their world knowledge, or impact their interpretation of context and meaning may increase image complexity. Regarding schema mismatch, P4 indicated,

> If you're placing something within a context that doesn't match the normal context that that person might understand that image, it becomes more difficult. If you have a banana, but it's, um, in some office and there's all like staplers and office supplies and computers and things all around, that sort of context, it maybe doesn't make as much sense as if it's within a kitchen setting or a dinner table, somewhere where you're able to make sense of it.

In a similar manner, six participants noted that context meaning can be obscured when elements, especially those in foreground of the image, are distracting or run counter to the overall meaning. Distorting an individual's world knowledge by capturing the photograph at a slanted angle, unusual vantage point, creating unnatural color contrasts (e.g., through poor color balance or high levels of saturation), or by using a lens that does not reflect how our eye naturally sees, such as a wide-angle lens, also dilutes the context. For instance, P7 noted,

> I actually think that because the color is a little extra saturated here. It does visually make the image even more complex. If you push the saturation too far, whereas letting it be a little bit more natural might actually help the eye move around.

Depth of field refers to the zone of the scene that remains in focus (Nagahara et al., 2008). Although depth of field may be a useful composition tool for guiding attention (see below), depth of fields that are too shallow may blur out and decrease the amount of information available to the viewer, making the image more about a singular object (e.g., a symbol) than depicting a clear routine or event in context. This point is illustrated clearly in the following comment by P2:

> It depends on the amount of blur. If I can still determine that what is around that orange are food items and a tray on a lunch table, then my imagination will make up the rest of the scene. So, I can identify that, okay, this is an orange on the table, but I'm still focused on the orange. The things around it still have context and give context to place for the orange. If it's slightly blurred, the orange stands out. If the orange is so blurred that we've lost the context, at that point it is truly about the orange and not the orange at lunch time.

Finally, as noted by one participant, a scenario with an overall unclear event of scenario becomes more complex for the viewer.

The Number and Completeness of Elements

A total of 11 participants discussed how the number of scene elements and shape may contribute to increasing image complexity. However, regarding the number of scene elements, participants discussed that, due to image schema, the viewer does not necessarily evaluate all of the physical items within a scene. Therefore, the actual number of physical elements in the scene may only contribute to complexity if the elements challenge one's schema, do not semantically support context, and/or draw attention away from primary focal elements though salient factors. For instance, P10 explains,

> If you had an image which was predominantly made up of full trees that have lots of green leaves and then in the midst of those trees is one red cardinal. You know, one brightly colored bird, which exhibits color contrast to the remainder of the image…we might say that that's a complex image because there's so much going on, you know, if you were trying to find the, the one leaf that was a different shape than the others, it would be, it'd be really difficult visually. However, you know that red bird, and all of a sudden, if you showed that image they might say, well, no, no, that's not a complex image at all, if we're trying to identify the bird, because there's this element of contrast.

When whole elements are not depicted, overlapping and cutoff/incomplete shapes may increase the abstractness of objects within the image, possibly decreasing an individual's ability to identify image elements. Furthermore, when these cutoff shapes/incomplete shapes extend beyond the border of an image, this phenomenon may also draw the individual out of the scene.

Age, Cognition, Environment, and Interests Can Potentially Compound Complexity

A total of five participants discussed that, in addition to more general factors that contribute to complexity, complexity itself may be compounded by a variety of person-centered factors, such as an individual's age, interests, environment, cognitive level, and culture, which need to be considered on an individual basis. For instance, as children develop visuospatial skills, understanding images taken at another vantage point may become easier to comprehend, or an individual with a gaming background may find packaging of a familiar game easier to understand than someone without experience in gaming. Furthermore, as P4 describes, how included elements support context may differ between individuals:

> Are you looking at from the American culture or looking at it from any other sort of country or ethnic background? Like, people would interact with the image differently, you know, maybe if maybe there are only fruits that are not even offered in certain places.

How Image Complexity May Impact the Viewer

Complexity Decreases Accessibility and Increases Difficulty in Understanding Image Content

A total of 12 participants discussed how image complexity may decrease the efficiency in which visual communication takes place through an image, making the image difficult to understand. This is well described by P4, who indicated,

> That's why decreasing the complexity is important in that, we're not thinking about the image, we're thinking about a communicative act, that the image is not about the image. The image is about communicating quickly and not having to decipher and translate what that thing might mean.

Therefore, participants discussed how image complexity may increase the effort required from the viewer to understand scene content, possibly causing increased visual load; increasing levels of fatigue, stress, frustration, confusion, and strain; and possibly decreasing contextual understanding. Based upon this added cognitive effort, four participants noted this issue may increase time for image understanding, with six participants noting that the viewer may eventually disengage with image content. For instance, paralleling a complex scene to "Where's Waldo?" images, P9 states, "And then eventually, you don't want to try anymore because Waldo, completely, to your knowledge is not in that picture, whether he's supposed to be or not." One individual also considered AAC access, noting that if objects are crowded together on the screen, a more efficient motor plan may be needed to select smaller communication targets.

Complexity May Be Mitigated When a User's Interests Are Activated

While the previous subtheme discussed the negative impacts image complexity may have on a viewer, five participants noted the broader impacts of how one's unique individual experience, motivation, and interests may impact how one responds when viewing a complex image. For instance, P10 noted, "If the image is aesthetically pleasing or something that is of interest to the, to the viewer, I think that would have a greater amount of impact than the complexity itself."

Composition Strategies to Decrease Image Complexity and Increase Cohesion

Composition Is Not a One-Size-Fits-All Approach and Principles Provide a Toolbox to Support Photographic Communication

All 13 participants discussed how compositional strategies described in the following subthemes, which aim to (a) decrease image complexity by helping draw attention to focal elements and (b) improve visual communication, may provide general guidelines for image creation. However, it is important to consider that the application of compositional strategies may need to be evaluated on an individual basis, depending upon the situation and communicative intent of the image. For instance, P7 said,

> So, I really think when it comes to organizing elements in an image, there's not necessarily one catch-all, you kind of have to have some awareness of the scene in front of you and then sort of have a toolkit at your disposal and picking the best tools for the scene.

Use of Contrast (Light and Dark)

A total of 13 participants discussed how contrast (light and dark) can be utilized to lower image complexity and decrease distraction. Bright objects draw our attention, and light-to-dark contrast creates a natural strong outline around the focal element(s). Therefore, having bright focal element(s) in the scene will help them stand out from the background. This purposeful application of light contrast is clearly illustrated by P11 through the following quote: "If we want to draw our attention to an item, we want that to be the brightest item in the scene."

Use of Contrast (Color)

A total of 12 participants discussed the use of color contrast in composition. Specifically, 12 participants discussed how scene element(s) of the most vivid color will attract attention. For instance, three participants discussed that warmer colors on a cooler background may draw our attention. It is also important to note that color is influenced by light, with colors that are well lit in the image becoming more vibrant. Therefore, light and color may work together in drawing attention to focal elements. When discussing color contrast, however, seven participants reported that naturally occurring color contrasts should be used, with P4 saying, "I would suggest keeping as many natural colors as possible but thinking about what colors you're photographing something on." However, participants indicated that utilizing bright and complementary color pairings (e.g., red and green, and those colors opposite each other on the color wheel) or using saturation to enhance color contrasts should be used cautiously to limit fatigue. Finally, two participants discussed grouping items by their natural internal color, indicating that grouping items by color may assist with building associations or provide balance, but the effects on limiting complexity are currently unclear.

Structural Principles

A total of 12 individuals discussed compositional principles related to structure, providing guidelines for balanced placement of key focal elements within the scene, including (a) the rule of thirds, (b) spiral line composition, (c) triangular composition, and (d) limiting items on the edge of the frame. First, placing focal items in central positions within the scene was discussed by 12 participants, indicating that central compositions are relatively simple, intuitive, and possibly better for smaller screen sizes. In addition to central composition, nine participants discussed the application of the rule of thirds and spiral line in creating a balanced composition. To apply a rule of thirds composition, the photographer breaks the image into thirds, both horizontally and vertically, to create a grid, which includes a total of nine small cells, similar to a tic-tac-toe board. The photographer then seeks to place focal elements of the scene on the intersecting grid lines (see Figure 2). The Fibonacci spiral line composition (see Figure 2) is a slightly more complex version of the rule of thirds and provides another method for creating images with natural balance. Spiral line composition can begin in any corner of the image and is based on a mathematical formula developed by Fibonacci that replicates a spiral line formation commonly seen in nature (e.g., a nautilus shell; Shuai, 2020). Together, participants discussed that the rule of thirds and spiral line composition may provide a guide for placing focal elements within the scene that may be more engaging than the central composition strategy, possibly supporting images with an increased number of objects, and perhaps better for larger screens. Regarding structural strategies, one of our experts stated,

> Some people are fond of the rule of thirds or the golden mean with, you know, kind of the Nautilus shell where that image is coming around. And we kind of focus where that primary third is. So, I mean, there are definitely compositional elements, or things that we can bring to help guide that eye or help bring that attention a little bit quicker into the frame.

An additional structural principle discussed by three participants involved arranging focal elements along implied triangular lines to help increase image structure (see Figure 3), though this may be less of an "on-the-go" strategy than the principles previously mentioned. Finally, to help avoid edge items that may increase image complexity and cause distraction, seven participants discussed keeping focal objects away from the edges of the frame, possibly leaving about 10%–20% of the image border clear.

Shape and Space

A total of 11 participants discussed the application of shape and space in limiting VSD complexity and decreasing distraction. Eight participants described how isolating focal element(s) by surrounding them with negative space can help draw attention to the object, limit object overlap and cutoff, support whole-object recognition, and decrease abstract shape formation. These points were illustrated by both P2 and P1, who stated,

> I can't emphasize enough, any time that you can create negative space around an object that is going to give preference to the object itself, and when a circular object is isolated, you know, in front of a plain background, it's very easy to understand what shape that is and identify that thing, but the more that it intersects with other shapes and overlaps, you know, the less simple it is.

However, one participant noted that, as individuals may be getting more used to using small-sized screens (e.g., cell phone), it is possible that individuals may be becoming more accustomed to identifying objects that are surrounded by limited space. In addition, two participants noted that scene viewers may be attracted to elements that have a contrasting shape to their surroundings.

Leading Lines Guiding Toward Focal Object and/or Into Frame

Eleven participants in total discussed the use of actual and implied lines to emphasize focal element(s). Eight participants described how strong horizontal, vertical, and diagonal lines, such as those formed by a tree branch, an individual's arm or an image frame may lead an individual's attention to focal element(s) within the scene. For instance, P10 stated: "You know, leading lines are huge, like what, you know what, what kind of a line, a road, a fence, a river, a tree branch. Your, your eye lock on to and follow to your, to your subject."

The use of an implied line was noted by six participants. Implied lines refer to lines that are not actually visible in the image, such as those described above, but which can guide our attention. For instance, implied lines may include an individual's eye gaze or direction of movement and attract an individual's attention to focal elements(s). Furthermore, implied lines move in the same direction that an individual reads. For instance, in English, an individual walking in the frame from left to right may help bring our attention into versus out of the scene.

Scale

A total of nine participants described the role of scale in drawing the viewers' attention to focal scene elements. In further detail, six participants indicated that the items of largest scale in the scene will attract attention first, unless the item is just providing context. This sentiment about the use of scale was elucidated by one of our experts, saying:

> Yeah, so in my head, there's a square, and a big river and the full river is included, and there's a tiny little bird. But if you had cut into the river and so now the river is just the corner of the photo instead of completely included in the photo, then that's the context in the background. And then the bird becomes the focus because that's what's fully, fully shown.

However, participants also noted that scale should not be manipulated unnaturally to draw one's attention. To support increased scale, five participants suggested placing focal elements in the foreground of the image. In addition, six participants discussed filling the photographic frame with the focal activity versus taking it at a distance to help ensure focal elements were given importance in the scene. However, caution may need to be given to ensure enough context is provided in which the activity is taking place.

Focus

Twelve participants in total identified the role of focus in composition, with seven participants indicating the focal element(s) should be in sharp focus to draw the viewers' attention. To the same effect, 11 participants described how a shallow depth of field can be utilized to blur background elements, making them out of focus, to help the viewer attend to elements in focus. The use of depth of field was articulately described by P11, denoting:

> If your object of focus is, say, in the foreground, you would keep that in focus, and your background may be blurred out, so that the object in focus is easier to pick out within that image.

However, when blurring out the background, care may need to be given to ensure the viewer can still understand image context, as described in Theme 1.

Facial Facial Features Should Be in Focus and Well Lit

A total of four participants discussed that the viewers' attention is naturally drawn to facial features. Consequently, facial features should be clear and well lit. This lighting technique may also help the viewer connect with the subject, as outlined by P13,

> You've got three shots of kids interacting with the AAC device. To me, the most interesting one is the kid in the middle, because I can really connect with this face. I get some eye light. I first look at his eyes.

Not Distorting Reality to Support Consistency With the Viewers' Image Schema and Support Interpretation of Meaning

A total of 11 participants discussed composition methods that may help support interpretation of meaning and help provide consistency between the image, the viewers' image schema, and their view of reality. Specifically, nine participants noted that the number of scene elements and background patterns or colors *may* need to be simplified *if* they do not contribute to meaning and/or are drawing focus away from focal element (e.g., nonfocal objects are bright and vivid in the scene). Furthermore, three participants noted the inclusion of familiar objects, locations, and meaningful engagements in supporting item recognition. For instance, P4 stated, "So, I'm thinking of instead of just having a general toy the person using the device and the person who's potentially helping them, to make an image of their own toy." Six participants also discussed how considering the perspective (e.g., height) of the individual may lower visuospatial demands. Regarding perspective, P13 discussed,

> So my tips would be, if this is for the child, then get on the child's level because that's their perspective, a lot of times, adults will stand up over them and they'll have this angle that's not conducive to the child's worldview and so they'll not connect with it.

Regarding limiting distortion within the image, four participants outlined how lenses of 50–85 mm, or equivalent, more naturally replicate how the eye naturally sees and may help limit line and depth distortions within the image. In addition, two participants indicated that taking photographs squarely versus at a tangential angle can help limit line distortion (e.g., helping straight lines look straight and not bent). Demonstrating this concept of taking focal elements in the scene squarely, P7 reports,

> If you are too off to the side of something or too above or below, like in relation to the scene that you're photographing that can sort of create angles that add complexity, rather than trying to square up more straight on.

Strategies to Support the Application of Composition Strategies in a Just-in-Time Setting

A total of seven participants discussed how grid lines and frame overlays are available on cameras to support the implementation of structural principles of composition. In this regard, seven participants discussed using "built-in" frame overlays, which provide a guide for implementing structural composition by the rule of thirds and spiral line. The availability and application of these guides were discussed by one expert, who said,

> There's certainly camera settings like in the point and shoot, the very consumer level. That allows you to put those (referring to rule of thirds and spiral line composition) kind of grids and it's invisible on the backside that helps kind of you compose. Yeah, yeah. Certainly, I know several cameras I have kind of puts those little grids in there for you.

Furthermore, two participants discussed using edge frame guides (e.g., a guide indicating the area 10%–20% from the frame edge) can help the photographer be aware of the frame border, or vignettes/border frames may help lead an individual into image content.

Automatic Camera Settings, Especially Aperture Priority

Seven participants in total discussed camera functions that automatically allow the photographer to set a specific exposure factor, such as film speed, shutter speed, or aperture, which controls depth of field. Once the photographer has set their primary exposure factor, the camera will automatically adjust other settings to provide an image with correct exposure. Regarding automatic exposure, participants primarily emphasized the application of the aperture priority setting. Participants described that, by setting the camera to automatic aperture priority, the photographer automatically achieves the desired depth of field, with other exposure factors being automatically adjusted accordingly.

Automatic and Touch Screen Apps for Focus, Light, Exposure, and Depth of Field

Similar to many camera phone applications, six participants discussed the use of automatic touch screen applications that allow the individual to identify the target element they want to be well lit and in focus. In addition, camera applications such as portrait mode may also be used to set depth of field. The use of touch screen applications was distinctly discussed by one visual artist, stating:

> Stick your finger on it, if it's your phone, or whatever it makes it in focus. The other thing that I think is maybe almost as important is that it must have light on it, so it's clearly readable, whereas the unimportant things can fall into shadow or darkness. And in fact, new iPhones do amazing things, like portrait mode and stuff is really incredible.

Furthermore, one participant noted the use of high dynamic range exposure, which commonly accompanies photographic applications. High dynamic range exposure settings are helpful for limiting the possibility that nonfocal items will be too bright and appear washed out within the scene. For instance, one participant described,

> You're photographing the building, and so the face of the building is in shadow and the sky around it's really bright. High dynamic range will sort of automatically take multiple exposures, one for sort of the exposure for the sky and one for the face of the building and digitally combine those so that you have detail in both the sky and the building. That can be useful.

4 and 4.6k Cameras and Post Image Editing

While these applications require processing of the image after capture, limiting their applicability to a just-in-time setting, four individuals noted the role of photograph editing in supporting the application of compositional principles after the photograph is taken. For instance, a range of image manipulations such as cropping image to fit structural principles, changing the image's focal point, and object removal may be achieved post editing. Furthermore, 4 and 4.6k cameras allow for detailed images, allowing for a range of image manipulations after taking the photograph; as one participant stated,

> You can shoot an entire scene using 4.6k and then go into postproduction…thinking about a child with you know in braces and a physical therapist trying to figure out what's hurting and how can we help…if you shot it in 4k then you could fix it in a computer and you could zoom in and you're not having to reshoot or stress a child out for example by trying to hold them steady to get a close up or to get whatever. You can just shoot it in 4k and then then then the physical therapist will just have to learn a little bit about Adobe premiere [laughter] to be able to really to learn how to zoom and pan.

Discussion

As with visual processing demands, there are clear principles that can inform decision making. People can be taught how to capture more powerful photos through principles photographers use to tell effective stories, depending on the situation and their objectives. If retrospectively people can make judgments about whether or not a photo is good, then a review of the principles used to draw such conclusions can help people take better photos in the future. The goal is not to take perfect photos but to take a moment to consider whether a photo will have the maximum desired impact.

Through this study, participants identified the following issues related to complexity and cohesion in photos: (a) factors that increase complexity and decrease cohesion, (b) impact on the viewer, (c) strategies to optimize, and (d) strategies to support quick application in a just-in-time setting. In the following sections, an overview of each theme is provided, along with recommendations for clinical practice (see Table 3), AAC development (see Table 4), and photographic illustrations demonstrating the application of compositional principles (see Figures 1–4).

Factors Increasing Image Complexity, Decreasing Cohesion, and Impacts on the Viewer

A VSD should quickly and intuitively represent an event so that communication around that event can take place. If the VSD is too complex, it can be difficult to find the desired item to express a message. Furthermore, there may be uncertainty about what hotspot might best convey an idea. Such complex scenes in VSD are best avoided. The current study supports some previous findings but also adds some caveats when a scene rather than a single image is involved. Based on study findings, a complex scene may have the following challenges: (a) It incorporates minimal or nonpurposeful use of composition saliency principles; (b) it lacks a unifying theme, context, or consistency; and (c) it lacks complete depiction of elements. High cohesion in some areas may mitigate challenges related to complexity. This is particularly important when not all aspects of a scene are under the photographer's direct control. The intersection of study findings helps provide more nuance to previous investigations aimed at investigating a single aspect such as number or arrangement of color in isolation. For example, our results support previous findings from O'Neil et al. (2019) that the number of objects in an image alone does not in itself guarantee increased complexity as the viewer already possessing an image schema/framework. Due to this schema, the viewer does not have to identify and interpret all scene elements to understand scene meaning. In fact, the inclusion of relevant elements may actually facilitate scene processing (Wilkinson et al., 2012). The nuance added by the current finding is that it is plausible that increasing the number of elements in an image may contribute to scene complexity *if* compositional considerations are not considered. For instance, if nonrelevant elements in the scene are the most salient (e.g., the brightest), these nonfocal elements may draw the viewers' attention, causing increased distraction. The consistency of objects in the scene is also a factor. Existing research indicates that inconsistent scene objects may draw attention and are harder to identify than expected items (Wu et al., 2014). Thus, if included element(s) are not semantically relevant to the scene or support context, these objects may cause distraction and contradict the

Table 3. Compositional recommendations for lowering visual scene display complexity and examples.

Subtheme	Figure(s)	Strategy for implementation
Use of contrast (light/dark and color)	2, 4	a. Attempt to make focal elements in the scene brighter than other scene elements and limit harsh contrasts by using soft lighting and automatic camera settings. b. Try to make focal items the most vivid color, using natural color contrasts. However, be cautious of color saturation and complimentary color contrasts.
Structural principles	2, 3, 4	a. Possibly consider central composition for more simplistic scenes and smaller screens. Also, consider rule of thirds or spiral composition guides to making simple images more engaging, support organization of an increased number of elements, and larger display screens. For images with increased time for composition, one may also wish to consider arranging items along implied triangular lines to increase structure and use overlay guides, as possible. b. Try to avoid items on the edge of the image to limit cutoff, using an edge guide as available.
Shape and space	2, 3, 4	a. Try to provide space around elements to increase object identification through whole shape.
Leading line	1, 2, 3, 4	b. Consider using real and implied lines (e.g., natural line, individuals implied gaze during task engagement, individual's movement direction) to point toward focal elements and draw the viewer into the scene/frame.
Scale	2, 4	a. Place the focal element(s) in the foreground, when possible, and consider filling the frame with the activity/focal objects, leaving some space for context.
Focus	2, 4	a. Try to ensure the focal elements are in focus. Consider using a shallow depth of field to blur nonfocal items but still provide some context. Use automatic camera settings, as available.
In focus and lit facial features	1	a. Try to have facial features well lit to support engagement.
Not distorting reality to support consistency with the viewer's image schema and support interpretation of meaning	2, 3	a. Try to support an individual's scene schema by (a) keeping items in the scene, especially those in the search area, semantically congruent and supporting scene context; (b) taking the image from the perspective (e.g., height) of the individual reading the image; (c) not taking images from an offset angle; and (d) including familiar objects, locations, and meaningful engagements. When possible, use a lens that best represents how the eye sees (e.g., 50–85 mm). b. Consider limiting/simplifying the number of objects, background patterns, and/or colors *only if* they are highly salient/drawing and individuals focus away from focal element(s) or do not support context.

viewer's image schema, increasing complexity. Therefore, while in general the number of scene elements alone does not appear to impact image complexity, when possible, some consideration to the type of element(s) included in the scene and how they are composed may be beneficial.

Study findings related to whole versus partial objects can also be considered in regard to scene composition. Light and Drager (2007) reported that children commonly depict concepts that are grounded in context, using whole shapes and rarely incorporating isolated body parts or events. Therefore, it may be helpful in a scene to consider the number of overlapping/incomplete shapes and their location within the image. If multiple elements are presented in an overlapping fashion, these partial objects may become difficult to interpret and create abstract shapes that are difficult to identify. Furthermore, if focal objects are cut off by the frame of the scene, such as a person is half in and half out of the photograph, the cutoff shape may draw the viewer out of the image, lowering the photograph's level of visual communication. There remains nuance when considering the interaction of objects and the overall scene, however. For example, edge-of-frame items that remain recognizable and contribute to context may impact complexity less than focal objects or incomplete objects that have become abstract due to cutoff. For instance, if the tip of an individual's hair is cut off by the frame, the scene still likely provides enough information for the viewer to recognize the individual and understand their context and role in the depicted activity. However, if the frame cuts off the depicted individual midface, they can become less recognizable, which may decrease image context and visual communication (see Figure 4).

Finally, our findings support the notion that a lack of focus or too much image blurring may decrease available

Table 4. Recommendations and considerations for augmentative and alternative communication development to facilitate visual scene display image capture.

Providing overlay options, which include rule of thirds, spiral line, and edge guides to help for the application of quick application of structural compositional principles.
Provide automatic settings, similar to many phone applications, that allow the photographer to quickly identify item(s) within the scene which should be (a) the lightest and (b) in focus, as well as allowing for manipulation of depth of field.
Provide automatic high dynamic range exposure options to decrease distracting areas of brightness and high contrast by providing an even exposure.
Provide intuitive applications for post capture photo editing for use as needed (e.g., light and color contrast, depth of field, structural composition), possibly including 4–4.6k options.
Use of lenses that best reflect how the eyes see nature, such as a 50- to 85-mm lens or equivalent.

Figure 1. Images of a child during a block stacking activity. Images of a child during a block stacking activity indicating positive (left) and negative (right) use of composition principles. Selected positive compositional principles (left) include (a) face being well lit, (b) task engagement, and (c) leading line created by blocks, arms and eye gaze leading to focal activity and moving from left to right (as illustrated by the black arrows). Selected negative compositional principles (right) include (a) face is not well lit with too much shadow, (b) the child is not task-engaged, and (c) line leading the viewer out of the frame, as illustrated by the white arrow. The positive image (left) may be achieved by moving the child slightly out of the chair shadow and turning on an extra light.

context, making the photograph parallel line–drawn symbols commonly seen on AAC devices where the image is about an object versus a contextual event. Therefore, while image blurring may help support attention to focal scene elements, too much blurring may increase cognitive demands for the individual using AAC (Thiessen et al., 2016).

Person-Centered Factors

Beyond saliency and meaning, it is important to consider image complexity; how the individual reacts to complex images may vary depending upon person-centered factors. For instance, one participant described that a video game cover may be complex for an individual without a background in gaming or understanding of game content. However, the image may be simple to understand if the individual has background experience with the image style (e.g., use of a range of vibrant colors, characters). In a similar manner, individuals viewing complex images may experience increased levels of effort, fatigue, and frustration associated with difficulty understanding the image. This increased level of effort may ultimately mean the individual becomes disengaged with the image. However, if the viewer is interested and motivated to engage with the image, these negative effects may be minimized, and engagement may remain. For instance, a "Where's Waldo?" image incorporates many of the discussed factors that makes an image complex. However, "Where's Waldo?" fans may ultimately remain engaged and receive enjoyment from viewing the complex images. Therefore, while this study presents an initial framework for understanding factors contributing to image complexity, discussed factors may impact the viewer differently.

Composition Strategies to Decrease Image Complexity and Increase Cohesion

Study findings provide a framework for understanding how composition strategies such as contrast, structure, shape and space, line, and focus may decrease image complexity and increase cohesion, along with providing considerations for supporting the viewer's image schema. These findings are consistent with prior multidisciplinary works identifying principles behind how to take a good photograph (e.g., Uzun et al., 2014), factors impacting visual search in an item array (Heuer & Hallowell, 2007), and current research on VSD design. However, our results extend compositional considerations to VSDs. In further detail, to create implied lines within the image that support attention to focal objects, study findings continue to support the use of human figures and faces within VSDs (Wilkinson & Light, 2011) that are task-engaged versus looking at the camera (Thiessen et al., 2014, 2016, 2017) or another direction leading out of the frame. In addition, as it is well documented that human figures draw our attention (Wu et al., 2014), facial features in images should be well lit. However, following an individual's gaze or point requires joint attention skills (Wilkinson et al., 2012). Therefore, the use of implied line in images for those with limited or emerging joint attention requires further consideration. In a similar manner, experts also discussed that lines flowing in the direction an individual reads (i.e., left to right for Americans) may help draw people into the scene, but the application of this principle to those with limited to emerging literacy skills is unclear.

For grid designs, another factor that may help the individual using AAC to identify their target element is grouping by internal color (e.g., Wilkinson et al., 2008) and use of space (Light, Wilkinson, et al., 2019). Scenes by nature include multiple items. However, paralleling the work of Light, Wilkinson, et al., 2019, providing space around scene objects may limit overlap, supporting depiction of whole shapes and possibly supporting VSD success. In regard to color, it is well known that color draws our attention in scenes (Wolfe & Horowitz, 2017). However, while not largely discussed by participants, how color grouping may impact visual scene requires further consideration as scene elements are generally not as isolated as those in a grid format. For instance, it is possible that multiple overlapping objects of the same shade and color may hinder object recognition by making it harder for the viewer to distinguish different whole objects.

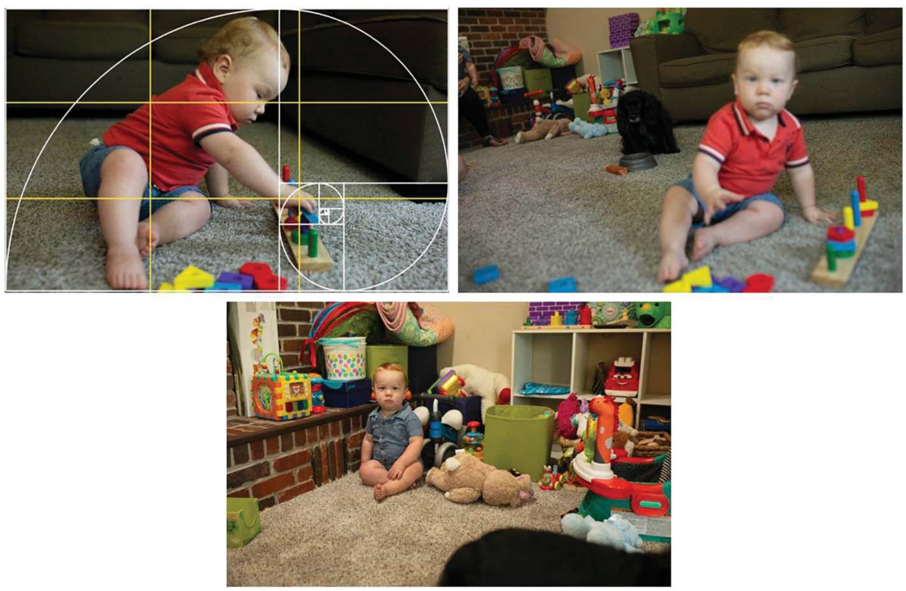

Figure 2. Images of a child during inside playtime. Images of a child during inside playtime, indicating positive (top left) and negative (top right and bottom middle) use of composition principles. Positive compositional principles (top left) include (a) task engagement; (b) child perspective; (c) being attracted to brighter focal elements; (d) spiral line composition (white spiral overlay), which also approximates rule of thirds (yellow grid overlay); and (e) largest objects drawing attention (scale), with the frame largely filled with the activity. Selected negative compositional principles for the top right image include (a) person and objects of interest are not in focus; (b) camera engagement; and (c) a nonsemantic element within the scene (a dog), which does not support the play context is in primary focus. Selected negative compositional principles for the bottom middle image include (a) camera engagement; (b) multiple overlapping shapes; (c) a nonsemantic object is in the foreground (a dog), which is not supporting context and is minimally recognizable due to cut off, leading out of the frame; (d) lack of image structure; and (e) a large amount of space around activity. The positive image (top left) may be achieved by waiting for activity engagement, ensuring the object and person of interest are in focus, and moving closer to the child to help fill more of the frame with the target activity.

In contrast, color grouping may help support an individual's learning of color associations and provide balanced images. Therefore, more research is needed into how internal color grouping impacts VSD complexity.

Avenues to support consistency with the viewer's image schema were also highlighted by our experts to support interpretation of meaning. Similar to McKelvey et al. (2010), study findings continue to support the use of familiar/personally relevant objects and meaningful engagements to help draw the viewers' attention to focal elements and support understanding of scene meaning. Furthermore, experts noted that taking the image from the perspective (e.g., height level) of the viewer may help the individual connect with the scene and lower visuospatial demands, possibly as objects may be more difficult to identify when shown from an unfamiliar perspective (Heuer & Hallowell, 2007). However, when capturing an image from the perspective of the viewer, existing research (e.g., Wilkinson et al., 2012) would indicate ensuring important people, including the individual with complex communication needs, remain included in the image, as appropriate. Furthermore, study findings highlight a range of structural principles such as central arrangement, rule of thirds, the Fibonacci spiral, and triangular arrangement, with the rule of thirds and the Fibonacci spiral possibly providing an engaging framework for arranging scene objects in comparison to centrally arranged displays. Structural principles as they relate to taking photos for VSD purposes are provided in Table 3, though further research is required to fully elucidate their role in VSD design for those with complex communication needs.

Clinical Applications and Considerations

VSDs can be captured in the moment to provide opportunities for communication. The findings from this study can be used to guide the models used for instructing parents, caregivers, and clinicians who are learning to use VSDs and guide future upgrades to help high-quality photos be taken with minimal effort. In terms of instruction, care should be taken not to overwhelm those new to making VSDs. Lengthy discussions about composition with parents and caregivers are not the goals of the current study. Instead, the models used for parents by educators, researchers, and technology developers should be carefully constructed to present maximally powerful images. Additionally, quick tips and examples, as shown through Table 3, can help individuals develop a photographer's eye for scenes. Furthermore, any premade VSDs used as examples or ways to get started should consider how to provide images with good composition strategies, as described. Clinical considerations for how compositional principles can be applied to help limit VSD complexity are outlined in Table 3, with study findings supporting existing practices for VSD design and providing new directions for clinical consideration and research.

Figure 3. Images of a child during outside playtime with mom. Images of a child during outside playtime with mom, indicating positive (top left) and negative (top right and bottom middle) use of composition principles. Selected positive compositional principles (top left) include (a) task engagement, (b) whole element depiction, (c) child perspective, and (d) use of triangular structure, as illustrated by the white arrows. Selected negative compositional principles for the top right image include (a) adult perspective; (b) cut of people on frame border; (c) line leading out of frame right to left, as illustrated by the white arrow; and (d) a nonsemantic element is included that does not contribute to context (a cell phone). However, in this instance, as the phone is not a highly salient feature, it may have limited impact on causing distraction. Selected negative compositional principles in the third image (bottom middle), include (a) eye gaze/line leading moving out of frame from right to left, as illustrated by the white arrow, and (b) a slanted angle distorting reality (i.e., they look like they are sitting on a hill). The positive image (top left) may be achieved by utilizing the child's perspective, waiting for task engagement, and taking the image squarely.

Considering compositional strategies as a toolbox for lowering complexity is especially important for capturing images in a just-in-time setting, and their application may help support just-in-time use of VSDs. For instance, Holyfield et al. (2019) recently provided guidelines for just-in-time programming VSDs for beginning communicators. In this article, their first three steps are to contextualize the interaction, capture engaging moments, and map vocabulary. Each of these important programming steps may be enhanced through considerations discussed in this article. Specifically, this study has outlined factors impacting context, and following, the authors further discuss strategies for promoting the quick application of compositional strategies for capturing important moments.

Regarding the application of compositional strategies, it should be considered that, in a just-in-time setting, the photographer will have minimal ability to manipulate scene content or consider a large number of compositional principles. However, it may be reassuring for someone to avoid worrying about cleaning up a room or the background when they can emphasize focal elements in a scene by minimizing any lines leading attention elsewhere. There may also be times when a parent or caregiver consults with a speech-language pathologist and there is more time to plan and design a VSD. In those cases, maximizing effectiveness could be worth a little extra time. Finally, if a team is noticing that a child is having difficulty using a VSD (i.e., appears to be staring at the display but not activating anything or is not engaging with the scene at all), then considering the composition of the display could be a useful troubleshooting tactic. Ultimately, the photographer should balance the feasibility of applying these principles, with available automatic camera supports, to capture the best image in a given situation.

For just-in-time application, the consideration of automatic camera settings and frame guides to "do the work" becomes increasingly important for capturing photographs that incorporate compositional principles. Our findings identified a number of avenues that may support the quick implementation of compositional strategies through automatic camera options. These automatic options included automatic frames (e.g., indicating the area close to the edge of the frame) and grid lines (for quick application of structural principles such as rule of thirds), along with automatic camera settings, such as (a) aperture priority for controlling image blurring/depth of field and (b) touch screen applications, similar to those available on cell phone cameras, which allow for quick adjustments in light, focus, and depth of field, along with helping provide more "even" exposures through high dynamic range exposures. Furthermore, participants discussed post image editing as an option to support photographic composition. Therefore, to allow for the application of these automatic photographic supports, AAC and software developers may wish to provide the identified automatic camera options to be available when capturing photographs for

Figure 4. Images of a child during pool play with mom. Images of a child during pool play with mom, indicating positive (left) and negative (right) use of composition principles. Selected positive compositional principles (left) include (a) vivid color on focal element and (b) a slightly blurred background that still provides enough information for context, (c) filling the frame with activity but leaving space for context, and (d) central composition. Selected negative compositional principles (right) include (a) similar colors (blue) in close proximity; (b) eye gaze leading out of image without task engagement, as illustrated by the white arrow; and (c) cutoff individual on the edge of frame. The positive image (left) may be achieved by choosing a different swimming outfit and ensuring the communication partner is more fully in the frame.

VSD use. Recommendations and considerations for AAC development are provided in Table 4. Finally, capturing cohesive scenes with a clear unifying theme may help programmers to match scene meanings to relevant vocabulary.

Limitations

Several limitations should be considered in regard to study findings. In more detail, the study included a limited sample size of composition experts, with only one having prior experience with individuals who have complex communication needs. Therefore, this lack of experience with those who have complex communication needs may mean their perspectives may be largely based on compositional impacts for neurotypical viewers and may have limited their ability to generalize their expertise. Finally, this investigation did not include the perspectives of those who use AAC and their support network, which is an important step in informing VSD design and use.

Future Directions

While the data provided through this investigation provide insights into photographic composition, further research is required to understand the impacts of composition principles on individuals with complex communication needs and explore the perspectives of individuals who use AAC and their support network. For instance, to establish the impact of identified compositional strategies, investigations utilizing quantitative methodologies such as reaction time (e.g., Thistle & Wilkinson, 2009) and eye-tracking techniques (e.g., Wilkinson & Light, 2011), especially during VSD use by those with complex communication needs, may help elucidate how compositional principles impact the viewers' attention and ability to identify focal elements. Utilization of VSD tasks is an important component to understanding the impacts of composition on VSD use by those with complex communication needs. The use of actual VSD-based tasks is important, as processing of salient (bottom-up) scene features may be influenced by the goal of the search task. For instance, Wolfe and Horowitz (2017) note that it is not task efficient to first attend to all the bright, colorful, and shiny objects in a scene if your task is to locate a cat. Consideration also must be given to differentiate how changes in individuals' visual attention patterns relate to improved comprehension of scene items. Furthermore, by considering how compositional strategies may be applied to increase scene cohesion, future research may consider how to best map scenes to corresponding vocabulary, especially for programming single words or phrases.

Considering display designs for AAC systems utilizing scenes, this study has only focused on static photographic images. However, it is plausible that the compositional considerations may also be applied to drawn symbol sets, such as those outlined by Worah et al. (2015), and video VSDs (e.g., Babb et al., 2020) to support visual communication for a range of image mediums. Furthermore, considering the AAC interface as a whole, factors such as the inclusion of text and a navigation bar may also contribute to the overall complexity of a VSD interface. For instance, placing the navigation bar at the top of the display in proximity to those included in the VSD may promote attention VSDs (O'Neil et al., 2019). Thus, whether the application of the discussed compositional factors may help provide guidance on design of the full VSD interface, which may include navigation bars, and text requires further consideration. For example, having the navigation bar increase in luminosity when attended to by the individual using AAC may help support attention via light contrast.

Thinking beyond commercial displays and access technologies, as AAC progresses to consider virtual and augmented reality, how the compositional strategies discussed within this study may be applied to 360° and frameless environments may provide guidelines for considering the design of new interfaces. Furthermore, some brain–computer interface-based methods for AAC (BCI-AAC) access require the individual to focus their attention on communication items in the AAC display (Brumberg et al., 2018). Therefore, how compositional factors may support an individual's attention to key scene elements during BCI-AAC presentation may help support outcomes. Finally, in parallel to scanning-based AAC access, some BCI-AAC techniques require communication items to be highlighted (Pitt, Brumberg, Bumison, et al., 2019; Pitt, Brumberg, & Pitt, 2019). Therefore, exploring how visual composition can be applied to support visual attention to highlighted scene elements may bolster VSD-related outcomes for scanning-based (McCarthy et al., 2018) and BCI-AAC techniques.

Acknowledgments

This work was supported by the Nebraska Tobacco Settlement Biomedical Research Development Fund. Furthermore, the authors would like to thank Megan Salley, Kathrin Sindt, and Rachel McMahon for their contributions to the project, along with Michaela Reddel for photography and Tony Pitt for image editing.

References

Babb, S., McNaughton, D., Light, J., Caron, J., Wydner, K., & Jung, S. (2020). Using AAC video visual scene displays to increase participation and communication within a volunteer activity for adolescents with complex communication needs. *Augmentative and Alternative Communication, 36*(1), 31–42. https://doi.org/10.1080/07434618.2020.1737966

Beukelman, D. R., & Light, J. (2020). *Augmentative and alternative communication: Supporting children and adults with complex communication needs* (5th ed.). Brookes.

Beukelman, D. R., Hux, K., Dietz, A., McKelvey, M., & Weissling, K. (2015). Using visual scene displays as communication support options for people with chronic, severe aphasia: A summary of AAC research and future research directions. *Augmentative and Alternative Communication, 31*(3), 234–245. https://doi.org/10.3109/07434618.2015.1052152

Brantlinger, E., Jimenez, R., Klingner, J., Pugach, M., & Richardson, V. (2005). Qualitative studies in special education. *Exceptional Children, 71*(2), 195–207. https://doi.org/10.1177/001440290507100205

Brown, J., & Thiessen, A. (2018). Using images with individuals with aphasia: Current research and clinical trends. *American Journal of Speech-Language Pathology, 27*(1S), 504–515. https://doi.org/10.1044/2017_AJSLP-16-0190

Brown, J., Thiessen, A., Beukelman, D., & Hux, K. (2015). Noun representation in AAC grid displays: Visual attention patterns of people with traumatic brain injury. *Augmentative and Alternative Communication, 31*(1), 15–26. https://doi.org/10.3109/07434618.2014.995224

Brown, J., Thiessen, A., Freeland, T., & Brewer, C. H. (2019). Visual processing patterns of adults with traumatic brain injury when viewing image-based grids and visual scenes. *Augmentative and Alternative Communication, 35*(3), 229–239. https://doi.org/10.1080/07434618.2019.1609578

Brumberg, J. S., Pitt, K. M., Mantie-Kozlowski, A., & Burnison, J. D. (2018). Brain–computer interfaces for augmentative and alternative communication: A tutorial. *American Journal of Speech-Language Pathology, 27*(1), 1–12. https://doi.org/10.1044/2017_AJSLP-16-0244

Creswell, J. (2012). *Qualitative inquiry and research design* (3rd ed.). Sage Publications.

Dietz, A., Weissling, K., Griffith, J., McKelvey, M., & Macke, D. (2014). The impact of interface design during an initial high-technology AAC experience: A collective case study of people with aphasia. *Augmentative and Alternative Communication, 30*(4), 314–328. https://doi.org/10.3109/07434618.2014.966207

Gibbs, G. R. (2008). *Analyzing qualitative data*. Sage Publications.

Hajjar, D. J., McCarthy, J. W., Benigno, J. P., & Chabot, J. (2016). "You get more than you give": Experiences of community partners in facilitating active recreation with individuals who have complex communication needs. *Augmentative and Alternative Communication, 32*(2), 131–142. https://doi.org/10.3109/07434618.2015.1136686

Henderson, J. M., & Hayes, T. R. (2018). Meaning guides attention in real-world scene images: Evidence from eye movements and meaning maps. *Journal of Vision, 18*(6), 10. https://doi.org/10.1167/18.6.10

Heuer, S., & Hallowell, B. (2007). An evaluation of multiple-choice test images for comprehension assessment in aphasia. *Aphasiology, 21*(9), 883–900. https://doi.org/10.1080/02687030600695194

Heuer, S., Ivanova, M. V., & Hallowell, B. (2017). More than the verbal stimulus matters: Visual attention in language assessment for people with aphasia using multiple-choice image displays. *Journal of Speech, Language, and Hearing Research, 60*(5), 1348–1361. https://doi.org/10.1044/2017_JSLHR-L-16-0087

Holyfield, C., Caron, J., & Light, J. (2019). Programing AAC just-in-time for beginning communicators: The process. *Augmentative and Alternative Communication, 35*(4), 309–318. https://doi.org/10.1080/07434618.2019.1686538

Hux, K., Buechter, M., Wallace, S., & Weissling, K. (2010). Using visual scene displays to create a shared communication space for a person with aphasia. *Aphasiology, 24*(5), 643–660. https://doi.org/10.1080/02687030902869299

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10–12), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203. https://doi.org/10.1038/35058500

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(11), 1254–1259. https://doi.org/10.1109/34.730558

Jagaroo, V., & Wilkinson, K. (2008). Further considerations of visual cognitive neuroscience in aided AAC: The potential role of motion perception systems in maximizing design display. *Augmentative and Alternative Communication, 24*(1), 29–42. https://doi.org/10.1080/07434610701390673

Light, J., & Drager, K. (2007). AAC technologies for young children with complex communication needs: State of the science and future research directions. *Augmentative and Alternative Communication, 23*(3), 204–216. https://doi.org/10.1080/07434610701553635

Light, J., McNaughton, D., & Caron, J. (2019). New and emerging AAC technology supports for children with complex communication needs and their communication partners: State of the science and future research directions. *Augmentative and Alternative Communication, 35*(1), 26–41. https://doi.org/10.1080/07434618.2018.1557251

Light, J., Wilkinson, K. M., Thiessen, A., Beukelman, D. R., & Fager, S. K. (2019). Designing effective AAC displays for individuals with developmental or acquired disabilities: State of the science and future research directions. *Augmentative and Alternative Communication, 35*(1), 42–55. https://doi.org/10.1080/07434618.2018.1558283

Liu, F., & Gleicher, M. (2006). Region enhanced scale-invariant saliency detection. In *Proceedings of the 2006 IEEE International Conference on Multimedia and Expo* (pp. 1477–1480). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/icme.2006.262821

McCarthy, J. W., Benigno, J. P., Broach, J., Boster, J. B., & Wright, B. M. (2018). Identification and drawing of early concepts in children with autism spectrum disorder and children without disability. *Augmentative and Alternative Communication, 34*(2), 155–165. https://doi.org/10.1080/07434618.2018.1457716

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282. https://doi.org/10.11613/bm.2012.031

McKelvey, M. L., Hux, K., Dietz, A., & Beukelman, D. R. (2010). Impact of personal relevance and contextualization on word–picture matching by people with aphasia. *American Journal of Speech-Language Pathology, 19*(1), 22–33. https://doi.org/10.1044/1058-0360(2009/08-0021)

Mukherjee, B., & Nair, A. G. (2012). Principles and practice of external digital photography in ophthalmology. *Indian Journal of Ophthalmology, 60*(2), 119–125. https://doi.org/10.4103/0301-4738.94053

Muraco, L. (2020). Improved medical photography: Key tips for creating images of lasting value. *JAMA Dermatology, 156*(2), 121–123. https://doi.org/10.1001/jamadermatol.2019.3849

Nagahara, H., Kuthirummal, S., Zhou, C., & Nayar, S. K. (2008). Flexible depth of field photography. In D. Forsyth, P. Torr, & A. Zisserman (Eds.), *Computer vision: Lecture notes in computer science* (Vol. 5305, pp. 60–73). Springer. https://doi.org/10.1007/978-3-540-88693-8_5

O'Neill, T., & Wilkinson, K. M. (2020). Preliminary investigation of the perspectives of parents of children with cerebral palsy on the supports, challenges, and realities of integrating augmentative and alternative communication into everyday life. *American Journal of Speech-Language Pathology, 29*(1), 238–254. https://doi.org/10.1044/2019_AJSLP-19-00103

O'Neill, T., Wilkinson, K. M., & Light, J. (2019). Preliminary investigation of visual attention to complex AAC visual scene displays in individuals with and without developmental disabilities. *Augmentative and Alternative Communication, 35*(3), 240–250. https://doi.org/10.1080/07434618.2019.1635643

Peterson, B. (2003). *Learning to see creatively: Design, color & composition in photography* (3rd ed.). Amphoto.

Pitt, K. M., Brumberg, J. S., Burnison, J. D., Mehta, J., & Kidwai, J. (2019). Behind the scenes of noninvasive brain–computer interfaces: A review of electroencephalography signals, how they are recorded, and why they matter. *Perspectives of the ASHA Special Interests Groups, 4*(6), 1622–1636. https://doi.org/10.1044/2019_PERS-19-00059

Pitt, K. M., Brumberg, J. S., & Pitt, A. R. (2019). Considering augmentative and alternative communication research for brain–computer interface practice. *Assistive Technology Outcomes and Benefits, 13*(1), 1–20. http://www.atia.org/atob

QSR International. (2018). *NVivo 12* (Version 12 plus) [Software]. https://www.qsrinternational.com/nvivo/products

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity, 52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Shuai, W. (2020). Mathematical thinking in photography. *International Journal of New Developments in Education, 2*(1), 36–68. https://doi.org/10.25236/IJNDE.2020.020109

Syed, M., & Nelson, S. C. (2015). Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood, 3*(6), 375–387. https://doi.org/10.1177/2167696815587648

Thiessen, A., Beukelman, D., Hux, K., & Longenecker, M. (2016). A comparison of the visual attention patterns of people with aphasia and adults without neurological conditions for camera-engaged and task-engaged visual scenes. *Journal of Speech, Language, and Hearing Research, 59*(2), 290–301. https://doi.org/10.1044/2015_JSLHR-L-14-0115

Thiessen, A., Beukelman, D., Ullman, C., & Longenecker, M. (2014). Measurement of the visual attention patterns of people with aphasia: A preliminary investigation of two types of human engagement in photographic images. *Augmentative and Alternative Communication, 30*(2), 120–129. https://doi.org/10.3109/07434618.2014.905798

Thiessen, A., Brown, J., Beukelman, D., & Hux, K. (2017). The effect of human engagement depicted in contextual photographs on the visual attention patterns of adults with traumatic brain injury. *Journal of Communication Disorders, 69,* 58–71. https://doi.org/10.1016/j.jcomdis.2017.07.001

Thistle, J. J. (2019). The effect of symbol background color on the speed of locating targets by adults without disabilities: Implications for augmentative and alternative communication display design. *Perspectives of the ASHA Special Interest Groups, 4*(6), 1482–1488. https://doi.org/10.1044/2019_persp-19-00017

Thistle, J. J., & Wilkinson, K. (2009). The effects of color cues on typically developing preschoolers' speed of locating a target line drawing: Implications for augmentative and alternative communication display design. *American Journal of Speech-Language Pathology, 18*(3), 231–240. https://doi.org/10.1044/1058-0360(2009/08-0029)

Thistle, J. J., & Wilkinson, K. (2017). Effects of background color and symbol arrangement cues on construction of multi-symbol messages by young children without disabilities: Implications for aided AAC design. *Augmentative and Alternative Communication, 33*(3), 160–169. https://doi.org/10.1080/07434618.2017.1336571

Uzun, M., Bülbül, M., Toker, S., Beksaç, B., & Kara, A. (2014). Medical photography: Principles for orthopedics. *Journal of Orthopaedic Surgery and Research, 9*(1), 23. https://doi.org/10.1186/1749-799x-9-23

Wilkinson, K. M., Carlin, M., & Thistle, J. (2008). The role of color cues in facilitating accurate and rapid location of aided symbols by children with and without Down syndrome. *American Journal of Speech-Language Pathology, 17*(2), 179–193. https://doi.org/10.1044/1058-0360(2008/018)

Wilkinson, K. M., & Jagaroo, V. (2004). Contributions of principles of visual cognitive science to AAC system display design. *Augmentative and Alternative Communication, 20*(3), 123–136. https://doi.org/10.1080/07434610410001699717

Wilkinson, K. M., & Light, J. (2011). Preliminary investigation of visual attention to human figures in photographs: Potential considerations for the design of aided AAC visual scene displays. *Journal of Speech, Language, and Hearing Research, 54*(6), 1644–1657. https://doi.org/10.1044/1092-4388(2011/10-0098)

Wilkinson, K. M., & Light, J. (2014). Preliminary study of gaze toward humans in photographs by individuals with autism, Down syndrome, or other intellectual disabilities: Implications for design of visual scene displays. *Augmentative and Alternative Communication, 30*(2), 130–146. https://doi.org/10.3109/07434618.2014.904434

Wilkinson, K. M., Light, J., & Drager, K. (2012). Considerations for the composition of visual scene displays: Potential contributions of information from visual and cognitive sciences. *Augmentative and Alternative Communication, 28*(3), 137–147. https://doi.org/10.3109/07434618.2012.704522

Wilkinson, K. M., & Snell, J. (2011). Facilitating children's ability to distinguish symbols for emotions: The effects of background color cues and spatial arrangement of symbols on accuracy and speed of search. *American Journal of Speech-Language Pathology, 20*(4), 288–301. https://doi.org/10.1044/1058-0360(2011/10-0065)

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour, 1*(3), 0058. https://doi.org/10.1038/s41562-017-0058

Worah, S., McNaughton, D., Light, J., & Benedek-Wood, E. (2015). A comparison of two approaches for representing AAC vocabulary for young children. *International Journal of Speech-Language Pathology, 17*(5), 460–469. https://doi.org/10.3109/17549507.2014.987817

Wu, C.-C., Wick, F. A., & Pomplun, M. (2014). Guidance of visual attention by semantic information in real-world scenes. *Frontiers in Psychology, 5*(54), 10. https://doi.org/10.3389/fpsyg.2014