

Old Dominion University

ODU Digital Commons

Communication Disorders & Special Education
Faculty Publications

Communication Disorders & Special Education

2021

Using Eye Tracking to Explore Visual Attention in Adolescents With Autism Spectrum Disorder

Anne M.P. Michalek

Old Dominion University, aperrott@odu.edu

Jonna Bobzien

Old Dominion University, JBobzien@odu.edu

Victor A. Lugo

Old Dominion University

Chung Hao Chen

Old Dominion University, cxchen@odu.edu

Ann Bruhn

Old Dominion University, abruhn@odu.edu

See next page for additional authors

Follow this and additional works at: https://digitalcommons.odu.edu/cdse_pubs



Part of the [Cognition and Perception Commons](#), [Dental Hygiene Commons](#), [Psychological Phenomena and Processes Commons](#), and the [Special Education and Teaching Commons](#)

Original Publication Citation

Michalek, A., M. P. , Bobzien, J., Lugo, V., A., Chen, C. H., Bruhn, A., & Giannakos, M. (2021). Using eye tracking to explore visual attention in adolescents with Autism Spectrum Disorder. *International Journal of Biomedical and Clinical Engineering*, 10(1), 18 pp. <https://doi.org/10.4018/IJBCE.2021010101>


This Article is brought to you for free and open access by the Communication Disorders & Special Education at ODU Digital Commons. It has been accepted for inclusion in Communication Disorders & Special Education Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

Authors

Anne M.P. Michalek, Jonna Bobzien, Victor A. Lugo, Chung Hao Chen, Ann Bruhn, and Michail Giannakos

Using Eye Tracking to Explore Visual Attention in Adolescents With Autism Spectrum Disorder

Anne M. P. Michalek, Old Dominion University, USA

 <https://orcid.org/0000-0001-6850-3948>

Jonna Bobzien, Old Dominion University, USA

Victor A. Lugo, Old Dominion University, USA

Chung Hao Chen, Old Dominion University, USA

Ann Bruhn, Old Dominion University, USA

Michail Giannakos, Old Dominion University, USA

Anne Michalek, Old Dominion University, USA

ABSTRACT

Video social stories are used to facilitate understanding of social situations for adolescents with autism spectrum disorder (ASD). This study explored the use of eye tracking technology to understand how adolescents with and without ASD visually attend to video social story content and whether visual attention is related to content comprehension. Six adolescents, with and without ASD, viewed a video social story of visiting a dental office. Eye gaze metrics, including fixation duration and count, and visit duration were collected to measure visual attention, and a knowledge assessment was administered for comprehension. Results indicated adolescents with ASD fixated and maintained visual attention at rates lower than peers without ASD. Adolescents with ASD scored higher ($M=77.78$) than peers without ASD ($M=72.22$) on the assessment indicating no relationship between eye gaze metrics and knowledge accuracy. Impact and implications of visual image type on frequency and duration of visual attention generated by participants is discussed.

KEYWORDS

ASD, Auditory Attention, Personalized Video Social Stories, Visual Attention

INTRODUCTION

Autism Spectrum Disorder (ASD) is a heterogeneous, neurodevelopmental disorder affecting approximately 1% of the population worldwide (Centers for Disease Control and Prevention [CDC], 2018). Behavioral manifestations of ASD include deficits associated with social-communication skills and pragmatics, as well as the presentation of stereotypical/repetitive behaviors, strict adherence to routines, and limited interests that can persist through the person's lifetime. Furthermore, individuals

DOI: 10.4018/IJBCE.2021010101

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

with ASD often display sensory processing differences resulting in hypo/hypersensitivity to environmental sensory information (American Psychiatric Association, 2013). The combination of poor social communication skills, sensory sensitivity, and desire for sameness experienced by individuals with ASD, often results in extreme anxiety and/or stress in unfamiliar or novel situations during which social interaction is required. For some individuals with ASD, this stress and/or anxiety may manifest as inappropriate self-stimulatory or stereotypical behaviors, such as screaming, hitting, or running, which further impedes participation and inclusion in social interaction activities. Because individuals with ASD have difficulty predicting what will occur during novel situations, it is essential to identify and utilize empirically based methods that assist these individuals to prepare for unfamiliar or novel situations in advance. Exposing individuals with ASD to the potential activities and/or interactions that may occur during a novel situation in advance may reduce the occurrence of inappropriate behaviors rooted in stress and/or anxiety triggered by the situation. As such, the individual with ASD may become more amenable to experiencing novel or unfamiliar situations requiring social interaction, thereby facilitating an increase in successful social outcomes.

For over 20 years, educators and practitioners have been designing interventions for children and adolescents with ASD using social stories (Delano & Snell, 2006; Gray & Garland, 1993; Kokina & Kern, 2010; Kuoch & Mirenda, 2003; Reynhout & Carter, 2006; Sansosti & Powell-Smith, 2006; Scattone, 2008). Originally developed in 1993 by Carol Gray, Social Stories™ are short books describing a social event from the perspective of the person reading the story (Gray 2004; Gray & Garland, 1993; Qi, Barton, Collier, Lin, & Montoya, 2018). Social stories are reviewed prior to a specific social situation and are not utilized to elicit a specific change in behavior, but to facilitate the understanding of social situations and the perspectives of other individuals in the setting (Gray, 2004). Traditional social stories often pair vocabulary and language, commensurate with the communication level of the child or adolescent, with a series of images or photographs that illustrate a situation, event, or expectation and provide descriptions of the roles others play in the sample situation. Furthermore, the social story provides the child or adolescent with ASD an example of the sequence of events and reflects environmental cues and social interactions likely to occur during the situation. Using social stories with children or adolescents who have ASD primes the use of appropriate social responses by providing a model of the skills needed to engage successfully in the situation (Delano & Snell, 2006; Sansosti & Powell-Smith, 2006; Scattone, 2008). As such, the use of social stories has resulted in positive outcomes such as reducing disruptive behaviors (Scattone, Wilczynski, Edwards, & Rabian, 2002); decreasing inappropriate social behaviors (Graetz, Mastropieri, & Scruggs, 2009), and increasing prosocial behaviors for preschool children with ASD (Crozier & Tincani, 2007). Additionally, social stories as an intervention are attractive to practitioners because of the ease of implementation, individualization, and modification (Bozkurt & Vuran, 2014).

Despite the positive outcomes reported with the use of social stories to reduce inappropriate and disruptive behaviors (Graetz et al., 2009; Scattone et al., 2002), the literature is inconsistent with regard to the impact of social stories on internal, self-regulatory processes. For example, Gray and colleagues (1993) reported the use of video social stories with children with ASD increased appropriate engagement during novel events requiring social interaction. However, it was unclear whether the use of video social stories led to positive outcomes because the intervention increased predictability in the novel setting, which, in turn decreased anxiety and/or stress, or because repeated exposure to the social story resulted in the learning of actual social skill(s) that generalized to the novel setting. Furthermore, although most social story interventions include embedded photographs of the child or adolescent targeted by the story, it is unclear whether the child or adolescent reading/watching the social story actually visualize themselves as the primary participant in the story, and consequently, the setting (Sansosti & Powell-Smith, 2006). Unfortunately, it can be extremely difficult to generate photographs or videos of a child or adolescent with ASD engaging in appropriate social interaction skills in a novel setting that can then be used to create a social story for that child or adolescent. Therefore, new technologies should be explored that help practitioners create personalized social stories.

Chen, Bobzien, Bruhn, Zhang, and Hsu (2014) sought to address the lack of individualization in social stories by creating personalized video social stories. This study (Chen et al., 2014) measured the impact of a video social story created using Video Face Replacement Technology (VFRT) on the occurrence of behaviors associated with anxiety and fear in children with ASD during routine dental examinations. VFRT incorporated the use of a computational method and multimedia application to transpose the child's lifelike facial features and skin tone onto a model appearing in a video social story about going to the dentist (Chen et al., 2014). Three children with ASD, between the ages of nine and ten, whose parents had reported extreme difficulty during dental visits as indicated by inappropriate behaviors displayed by the child during the dental visit (e.g. screaming, kicking, biting, crying), participated in the study (Chen et al., 2014). Responses given by the participant's parents on a dental fear survey adapted from the Child Fear Survey Schedule-Dental Subscale (Cuthbert & Melamed, 1962) prior to the start of intervention indicated all participants with ASD appeared *very afraid/anxious* during dental visits. During intervention, each participant watched a personalized video social story following a decreasing weekly viewing schedule (e.g. five times during week one, four times during week two, three times during week three, etc.) prior to visiting the dentist. As such, each participant with ASD had the opportunity to view "themselves" engaging in specific, appropriate behaviors during the dental visit. Furthermore, the participants virtually experienced the sequence of activities followed during the dental visit (Chen et al., 2014). Following intervention, each participant visited the dental clinic depicted in the video social story for a cleaning and the participant's parents repeated the dental fear survey, this time focusing on the child's behaviors during this dental cleaning appointment. Observational data results indicated all participants demonstrated few inappropriate behaviors and more importantly, all participants imitated several phrases and appropriate behaviors modeled in the video social story during the dental cleaning appointment. Likewise, parents perceived a considerable reduction in fear/anxiety related behaviors during the follow-up dental appointment. Despite these positive results, data collection techniques utilized did not indicate whether the individualization and personalization of the video social stories actually contributed to positive outcomes or whether the positive outcomes were simply a result of the use of the video social story intervention itself.

Subsequently, Chen and colleagues (2015) sought to extend the research on video social story design, implementation, and content by measuring the impact of personalized image quality and information saliency on attention and processing of adolescents with ASD. Using a video learning analytics system (VLAS; Giannakos, Chorianopoulos, & Chrisochoides, 2015), Chen and associates (2015) tracked and quantitatively calculated the participant's interaction (e.g. accessing the play, pause, backward and/or forward controls) with two digitally rendered social story videos of going to the dentist. Although both video social stories were generated using the same digital template, one of the video social stories incorporated photographs of a generic peer model while the second incorporated VFRT to create a personalized video social story (Chen et al., 2014; Chen et al., 2015). Results indicated that both participants with ASD attended more consistently, and interacted more often, with the personalized video social story when compared to the video social story with the generic peer model (Chen et al., 2015). Once again, although increased attention and interaction were demonstrated during the use of a personalized video social story, it was unclear which aspects of the visual content within the personalized social story encouraged the increased visual attention.

In an effort to ascertain whether children and adolescents with ASD demonstrate a preference to specific types of visual content, there has been an increase in research using eye tracking technology as a data collection metric, especially as eye gaze relates to socially relevant visual stimuli (Guillon, Hadjikhani, Baduel, & Roge, 2014; McPartland, Webb, Keehn, & Dawson, 2011; Mondloch, Grand, & Maurer, 2002; Mondloch, Robinson, Maurer, 2010; Renzi, Shiavi, Carbon, 2013; Tsang, 2018). Eye tracking metrics reflect underlying cognitive processes by determining the allocation of visual attention (e.g. where a person is looking) when viewing an image, an object, or a person (Karatekin, 2004; Karatekin, 2007; Karatekin, Marcus, & Couperus, 2007). As such, research incorporating eye

tracking data collection metrics has the potential to provide answers to relevant questions related to clinical interventions and individual differences demonstrated by people with neurodevelopmental disorders (Venker & Kover, 2015). To date, research specific to children and adolescents with ASD that utilized eye tracking data collection has focused on facial emotion recognition (Tsang, 2018), face region preferences during social exchanges (Mondloch et al., 2002, 2010; Renzi et al., 2018), and patterns of visual attention which reflect social function and/or deficits (McPartland, Webb, Keehn, & Dawson, 2011). Although eye tracking data collection metrics have been used to explore visual attention to human faces by children and adolescents with ASD, there is paucity of research using this metric to determine if specific visual content/stimuli results in variability in visual attention to video social stories.

As such, the purpose of this research is to determine the feasibility using eye tracking metrics and a content comprehension measure to identify differences in visual attention during personalized video social stories by adolescents with and without ASD. Specifically, this study seeks to identify the relationship between specific visual content in personalized social stories and the comprehension and internalization of the social story content.

METHODS

First, we identified three categories of visually salient social story content: object, person, and environment. Then, we engaged adolescents with and without ASD in watching personalized social story videos. We had both groups answer multiple choice questions related to social story knowledge which aligned with object, person, environment, and information presented auditorily only. Finally, we analyzed group differences to delineate relationships between personalized social story video content and visual attention processes.

Participants

Six adolescents, three diagnosed with ASD and three without a diagnosis of ASD participated in the study. All six participants ranged in age from 11 – 16 years, all participated in general education classrooms at the elementary or secondary level, and all displayed visual acuity within normal limits according to parent report. Participant 1 was a 16-year-old male diagnosed with high-functioning ASD. According to assessment reports provided by the parents, participant 1 displayed average intelligence and age-appropriate communication skills; however, he exhibited social skill differences (e.g. fleeting eye contact, little joint attention, etc.) and sensory aversion to tactile and olfactory stimulation. Participant 2 was a 14-year-old male also diagnosed with high-functioning ASD. Assessment reports provided by the parents of participant 2 indicated average intelligence, but deficits in both age-appropriate communication and socialization skills. Furthermore, participant 2 displayed sensory aversion to tactile and auditory stimulation. Participant 3 was a 13-year, 4-month old male who received a diagnosis of high-functioning ASD until he was approximately 6-years-old. Information provided on assessment reports supplied by his mother, denoted average intelligence and communication skills; however, he would display mild inappropriate behaviors (e.g. refusal, elopement, jumping, shouting) when anxious and/or frustrated. Participant 4 was a 15-year-old female whose parent reported had no previous, or current, diagnosis of a neurodevelopment and/or cognitive disability. Similarly, participant 5 was a 12-year-old female who was also considered to be both age-appropriate and developmentally appropriate in all areas when compared to same-age peers. Participant 6 was a 11-year old female who had not history of a neurodevelopmental and/or cognitive disability and was judged to be developmentally age-appropriate in all areas. Table 1 summarizes participant characteristics.

Table 1. Characteristics of participants

Participant	Gender	Age (years)	Diagnosed Disability
1	Male	16	High-functioning ASD
2	Male	14	High-functioning ASD
3	Male	13	High-Functioning ASD
4	Female	15	None
5	Female	12	None
6	Female	11	None

Materials

Six, three minute long video social stories depicting a visit to a dental clinic were generated using VFRT. Although all video social stories were generated from the same template, each story was personalized by including digitally rendered images of the corresponding participant. In order to render a high quality video, an approach created by Bitouk and colleagues (2008) was applied to measure the head position, resolution, color and lighting of the subject, which are used to provide guidance to seamlessly blend the target of interest into the image template. Next, three Gestalt properties (i.e., symmetry, continuity and proximity) were applied to provide feedback on the blending quality, which is summarized as the following.

Let Ω be the domain of definition of a scene image, $\Omega = R_B \cup R_S$. R_B and R_S denote the regions that belong to backgrounds and foregrounds, respectively. After the background identification, we know that most of the structured objects in the image are contained in a sub-region $R_S \subset \Omega$. Let P_o be the initial partition of Ω from a bottom-up segmentation method. Let a denote a uniform patch from the initial partition P_o . For $\forall (a \in P_o) \wedge (a \in R_S)$, a is one of the constituent parts of an unknown structured object. Based on the initial part a , we want to find the maximum region $R_a \subset R_S$ so that, the initial part $a \in R_a$ and for any uniform patch i , where $(i \in P_o) \wedge (i \in R_a)$, i should have some special structural relations that obey the non-accidentalness principle with the remaining patches in R_a . This is formulated as follows:

$$R_a = \arg \min_R (E[\partial R]) \text{ with } (a \in R) \wedge (R \in R_S) \quad (1)$$

where R is a region in R_S , ∂R is the boundary of R , $E[\partial R]$ is a boundary energy function, which is defined as follows [15]:

$$E[\partial R] = \frac{-\int \int_R f(x,y) dx dy}{L(\partial R)} \quad (2)$$

where $L(\partial R)$ is the length of the boundary of R . $f(x,y)$ is the weight function in region R . The boundary energy function provides a tool for measuring how good a region is. The criterion of region goodness depends on how the weight function $f(x,y)$ is defined. One can use $f(x,y)$ to encode any information (e.g., color or texture) over the region R .

Video Social Story Knowledge Measure

The research team collaborated on video social story content to identify visual and auditory cues, which signal content saliency and reflect comprehension. Those cues were labeled as object, person,

environment, or auditory, and multiple choice questions were developed targeting each cue. Three questions were developed per cue level for a total of 12 multiple choice questions. Questions related to object, person, and environment were designed so that they could only be answered if visual attention was given to the target cues. The multiple choice questions were added to the test protocol as the video social story knowledge measure. Participants were able to select one of the multiple choice answers using mouse clicks as each multiple choice question was individually presented on the computer screen following the conclusion of the video social story. Aligning eye gaze metrics with a comprehension measure provides insight regarding both gathered information and information generation (Serrano, Owens, & Hallowell, 2018).

Eye Tracking Technology

Eye tracking data were recorded and analyzed using the Tobii Pro X2-60 computer screen-based eye tracker with Tobii Studio analysis software. The Tobii Pro X2-60 records eye movements using infrared corneal reflective technology at a sampling rate of 60 Hz (i.e. approximately once every 16.23 milliseconds). Gaze data accuracy was within .4 degrees of visual angle and precision was within .34 degrees of visual angle. Tobii's eye tracking technology is effective for generating reliable and valid brain/behavior outcomes for children and adolescents (Melinder, Konijnenberg, & Sarfi, 2013; Richmond & Nelson, 2009). The electronic, personalized social story videos were saved in Windows Media Audio/Video File (.wmv) format and added directly to Tobii Studio as a movie element. The video social story played as a movie on a Dell UltraSharp U-12HM monitor with a resolution of 1920 x 1080.

In order to correspond to the video social story knowledge measure, the digital story was analyzed to isolate scenes containing relevant object, person, or environmental cues relating to each of the 12 multiple choice questions on the measure. Since the personalized video social stories contained dynamic visual stimuli, static scenes were created using Tobii studio analysis software. A total of 24 static scenes were created per social story. These scenes were approximately six seconds in length and identical across participants. Scene use facilitates the reliable and valid creation of areas of interest (AOIs) (Figure 1). AOIs were created by overlaying different shaped boxes on defined areas of visual stimuli which were of interest. Three different levels of AOIs were created: object, person, and environment and included in scenes, as appropriate. AOIs were consistent in size and shape across scenes and participants. Data from object, person, and environment AOIs were aggregated by the grouping function in Tobii Studio. An AOI group collects eye gaze metrics from single AOIs across the media element.

From the aggregated object, person, and environment AOIs, three dependent eye gaze metrics were calculated: (1) total fixation duration – reflects the amount of time the participant's gaze is maintained or is relatively still on an AOI; (2) fixation count – reflects the number of times the participant maintains gaze in an AOI; (3) total visit duration – visits are dwell times or the amount of time from when the participant first focuses on an AOI until he/she looks away. Eye gaze metrics, such as fixations and visits, have been used in other research projects and empirically confirmed as valid indices of visual attention (Boivin, Weiss, Chhaya, Seffren, Awada, & Sikorskii; Heuer & Hallowell, 2009).

Design and Data Collection Procedures

This exploratory study used a quasi-experimental, mixed methods group research design. Specifically, a 2 X 3 mixed design was used with visual cue type as the within subject manipulation and diagnosis (i.e. ASD vs. non-ASD) as the between subject manipulation. Two primary means of data collection were utilized, a video social story knowledge measure, and eye-tracking technology. Data were collected during individual participant sessions in an eye tracking lab located on the campus of an urban university located on the southeast coast of the United States.

Figure 1. Scene use facilitates the reliable and valid creation of areas of interest (AOIs)



(a)



(b)

Procedures

Upon arrival at the university, participants were introduced to the researcher who engaged in social conversation to develop rapport and help ease anxiety. After approximately five-minutes of initial, social conversation, participants were led to the eye-tracking lab. The researcher briefly explained the eye tracking technology and equipment. Then, the participants were seated approximately 25 inches from the computer monitor and encouraged to seek a comfortable resting position. Each participant was asked to watch calibration circles as they moved across the screen. Once the eye tracking equipment had been successfully calibrated, the researcher provided specific instructions regarding the experimental tasks to the participant. Specifically, that he/she would be watching a personalized video about going to the dentist. Participants were instructed to watch and listen to the video as best they could because they would be answering several multiple choice questions pertaining to the video after it was over. The researcher attempted to confirm understanding prior to starting the video by giving the participant an opportunity to ask questions. Once the participant indicated understanding of the process, the viewing session began. Each video was approximately three minutes in length and the social story knowledge measure appeared on the computer screen immediately following the end of the personalized video social story. The entire testing session took approximately 30 minutes from beginning to end, including researcher introduction and explanation, system calibration, video social story viewing, and the completion of the content measure

RESULTS

During all phases of this exploratory study, no participants questioned the process nor did any participants ask for clarification. As such, participants did not appear to have difficulty following

the directions as explained and on average, completed the video social storying viewing and video social story knowledge measure in approximately six minutes.

According to the described methods, social stories were analyzed, segmented, and coded to isolate salient features of object, person, and environment. Furthermore, the scenes, AOIs, and AOI groups were created according the methods describe above. Participant results were coded using a number and stored in Tobii Studio. Data were extracted and analyzed using the statistics function. Means for each dependent variable (i.e. three eye gaze metrics) and raw scores for the accuracy measure were generated and entered in SPSS. A two-way mixed or split-plot ANOVA was run for each of the dependent variables: total fixation duration, fixation count, total visit duration, and social story knowledge accuracy. Mean scores and standard deviations (SDs) are outlined according to group and salient feature for each of the dependent variables in Tables 2, 3, and 4. The mixed design included diagnosis of ASD as a grouping variable with two levels (i.e. ASD and no ASD) and visual cue feature as the second grouping variable with three levels (i.e. object, person, environment). Then, bivariate correlations were examined to identify possible relationships between visual attention to visual cue features and social story knowledge accuracy.

Total Fixation Duration

Total fixation duration was measured for each visual cue feature (i.e. object, person, and environment) by aggregating object, person, and environment areas of interest across the personalized video social story. Total fixation duration reflects maintenance of eye gaze on a certain location and was calculated

Table 2. Mean fixation duration for each visual cue type in adolescents with and without ASD

Group	N	Object		Person		Environment	
		M	SD	M	SD	M	SD
ASD	3	0.466	0.174	0.270	0.311	0.343	0.099
Non-ASD	3	1.667	0.564	3.607	0.779	0.660	0.274
Total	6	1.066	0.756	1.938	1.903	0.502	0.253

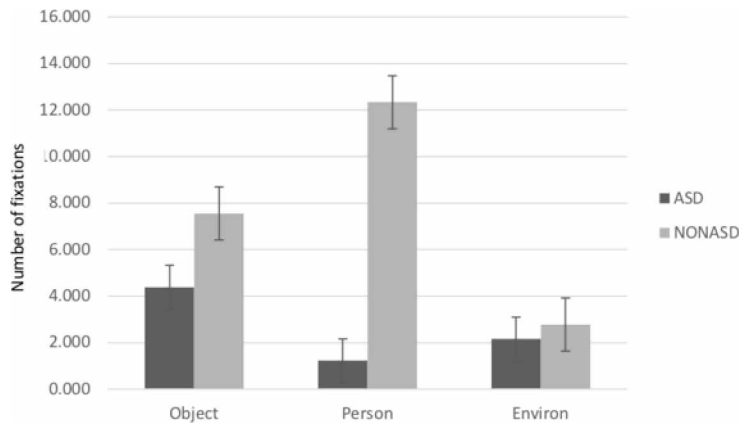
Table 3. Mean fixation count for each visual cue type in adolescents with and without ASD

Group	N	Object		Person		Environment	
		M	SD	M	SD	M	SD
ASD	3	4.389	2.699	1.222	1.347	2.156	0.583
Non-ASD	3	7.556	2.988	12.333	3.214	2.778	0.509
Total	6	5.972	3.082	6.778	6.473	2.467	0.596

Table 4. Mean visit duration for each visual cue type in adolescents with and without ASD

Group	N	Object		Person		Environment	
		M	SD	M	SD	M	SD
ASD	3	0.878	0.401	0.290	0.344	0.397	0.131
Non-ASD	3	1.919	0.610	4.051	0.816	0.690	0.267
Total	6	1.399	0.733	2.170	2.135	0.544	0.247

Figure 2. Mean fixation duration, reported in seconds, for object, person, and environment by diagnosis



in seconds. Between and within group differences was analyzed using a repeated measures ANOVA. Results indicated a main effect of visual cue $F(2, 8) = 17.106, p < .005$, eta squared = .810 and a main effect of group $F(1, 4) = 122.66, p < .005$, eta squared = .968. These main effects were subsumed under a two-way interaction effect of diagnosis and visual cue $F(2, 8) = 19.674, p < .005$, eta squared = .831. Figure 2 shows means and standard errors of mean fixation duration on object, person, and environment by diagnosis group.

In an attempt to specifically identify the effect of visual cue on between group results, a univariate ANOVA for object, person, and environment was performed. Group differences were noted for each of the visual cue types: object $F(1, 4) = 47.499, p < .005$, eta squared = .922; environment $F(1, 4) = 3.554, p > .005$, eta squared = .470; person $F(1, 4) = 47.499, p < .005$, eta squared = .922.

Fixation Count

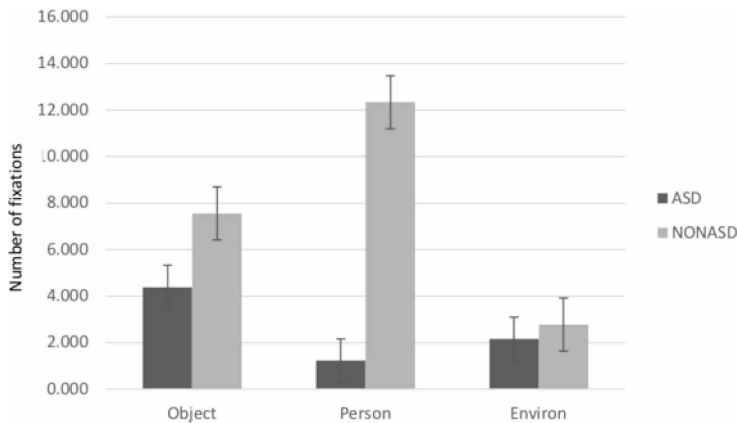
Fixation count was measured for each visual cue feature (i.e. object, person, and environment) by aggregating object, person, and environment AOIs across the personalized video social story. Fixation count measures the number of times the person fixates on an area of interest group. Using a repeated measure ANOVA, between and within group differences were analyzed. Resulted indicated a main effect of visual cue $F(2, 8) = 9.362, p < .05$, eta squared = .701 and a main effect of group $F(1, 4) = 59.787, p < .005$, eta squared = .937. These main effects were subsumed under a two-way interaction effect of diagnosis and visual cue $F(2, 8) = 13.336, p < .005$, eta squared = .769. Figure 3 shows means and standard errors of mean fixation count for object, person, and environment by diagnosis group.

Again, to identify the effect of visual cue specifically on between group results, a univariate ANOVA for object, person, and environment was performed. Group differences for person $F(1, 4) = 30.488, p = .005$, eta squared = .884 were demonstrated; however, no differences were noted between groups for object $F(1, 4) = 1.856, p > .05$, eta squared = .317 and environment $F(1, 4) = 1.941, p > .05$, eta squared = .327.

Total Visit Duration

Total visit duration for each visual cue feature (i.e. object, person, and environment) by aggregating object, person, and environment areas of interest groups were measured across the personalized video social story. A visit was defined as the amount of time, calculated in seconds, between the first fixation in the AOI group and the end of the last fixation in the AOI group with no fixations made outside the AOI. Total visit duration was calculated by adding the total number of visits made to an AOI group. A repeated measure ANOVA was utilized to analyze between and within group

Figure 3. Mean fixation count for object, person, and environment by diagnosis



differences. Results indicated a main effect of visual cue $F(2, 8) = 17.947, p < .005$, eta squared = .818 and a main effect of group $F(1, 4) = 130.246, p < .001$, eta squared = .970. These main effects were subsumed under a two way interaction effect of diagnosis and visual cue $F(2, 8) = 22.575, p < .005$, eta squared = .849. Figure 4 shows means and standard errors of mean total visit duration for object, person, and environment by diagnosis group.

A univariate ANOVA for object, person, and environment was used to specifically identify the effect of visual cue on between group results. Although group differences for person $F(1, 4) = 54.092, p < .002$, eta squared = .931 were noted, there were no differences between groups for object $F(1, 4) = 6.101, p > .05$, eta squared = .604 or environment $F(1, 4) = 2.918, p > .05$, eta squared = .422.

Video Social Story Knowledge Accuracy and Relationships Between Total Fixation Duration, Fixation Count, and Total Visit Duration

Tables 5 and 6 list the accuracy percentages, means, and standard errors for the participants with ASD and participants without ASD on video social story knowledge accuracy according to visual cue feature. Between group and within group differences were analyzed using a repeated measures ANOVA. A main effect of cue type $F(3, 12) = 2.100, p > .05$, eta squared = .344 was not indicated; however, a main effect of diagnosis $F(1, 4) = 145.800, p < .005$, eta squared .973 was noted.

Figure 4. Mean visit duration, reported in seconds, for object, person, and environment by diagnosis

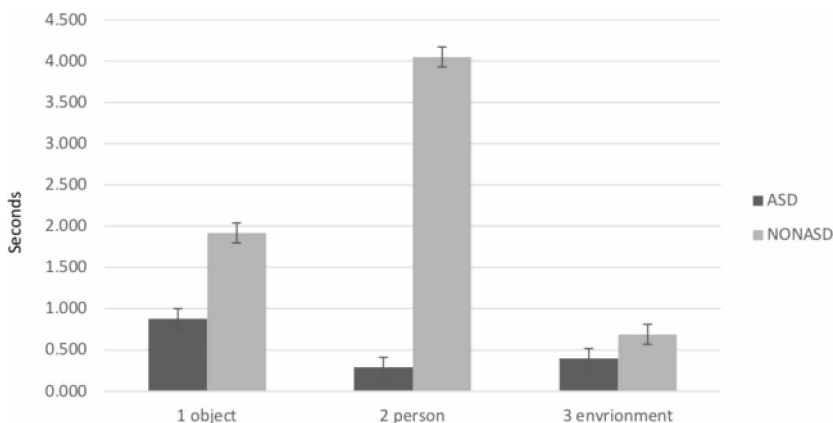


Table 5. Mean score on the social story knowledge measure for each visual cue type in adolescents with and without ASD

Group	N	Object		Person		Environment		Auditory	
		M	SD	M	SD	M	SD	M	SD
ASD	3	2.333	1.155	2.333	0.577	2.000	1.000	2.667	0.577
Non-ASD	3	2.667	0.577	2.000	1.000	1.333	0.577	2.667	0.577
Total	6	2.500	0.837	2.167	0.753	1.667	0.817	2.667	0.516

Table 6. Individual scores reported as percent accurate s on the entire social story knowledge measure for each adolescent with and without ASD

Participant Number	Percent Accurate
1	100%
2	58.33%
3	75%
4	75%
5	75%
6	66.67%

Furthermore, an interaction between diagnosis and cue type $F(3, 12) = .500, p > .05$, eta squared = .111 was not demonstrated. Finally, we wanted to determine if there was a relationship between eye gaze metrics and social story knowledge accuracy. A correlational analyses was conducted for adolescents with and without ASD between total fixation duration, fixation count, and total visit duration and social story knowledge accuracy on object, person, and environment questions. No relationships were found between total fixation duration, fixation count, and total visit duration and social story knowledge accuracy.

DISCUSSION

The purpose of this exploratory study was to determine if eye tracking metrics could be used to identify differences in visual attention for adolescents with and without ASD when watching a personalized video social story. Furthermore, the researchers sought to ascertain whether a relationship existed between patterns of visual attention and video social story knowledge accuracy. Eye tracking metrics were analyzed in relation to social story attributes and the participant’s ability to generate and maintain meaningful representations of social story content. Specifically, statistical analyses were conducted for adolescents with and without ASD on visual content organized as: object, person, or environment and social story knowledge accuracy questions for total fixation duration, fixation count, total visit duration. Although no relationship was indicated between total fixation duration, fixation count, and total visit duration and social story knowledge accuracy, differences in eye tracking patterns while viewing personalized video social stories were noted between adolescents with and without ASD.

Specifically, differences were noted across the three relevant eye gaze metrics measuring visual attention for the three relevant visual cue types: object, person, and environment.

Object

The personalized video social stories contained three objects (e.g. dentist chair, dentist tray, book participant was reading in the waiting room) identified as potential areas of interest. These objects were judged to be relevant to the theme of the story and potentially capable of increasing the participant's ability to internalize events and interactions. Overall, the adolescents with ASD did not fixate or sustain visual attention to the three target objects in the social story for the same length of time as the adolescents without ASD. However, no difference was noted in the frequency of visual attention to the target objects or in the length of time between when the adolescent initially looked at the object and when the adolescent looked away from the object. These results suggest that adolescents with ASD are scanning and adjusting eye gaze continuously; thereby, interrupting direct sustainment of visual attention. Despite a decreased amount of total fixation duration on salient objects, the adolescents with ASD are recognizing and paying visual attention to important, salient features. These results are consistent with Keehn and colleagues' (2009) indicating that a lack of sustained visual attention to object cues does not influence the ability to generate and retain knowledge about that object for individuals with ASD. As such, it appears that, regardless of the significant differences in visual attention patterns, adolescents with and without ASD are able to process object stimuli regardless of sustained visual attention.

Person

Images of three people within the personalized video social story were targeted as important areas of interest. Of these people, two were images of the participant and one was an image of an important character encountered during a visit to the dentist (i.e. the dental hygienist). Results indicated that adolescents with ASD do not visually attend to people for the same amount of time or in the same manner as adolescents without ASD. Differences in visual attention to persons in the video social story persisted across all three eye gaze metrics: total fixation duration, fixation count, and total visit duration. Visual analysis of the resulting heat maps indicated that adolescents with ASD spent less time looking at the facial areas of the person in the image and visually fixated on the image of themselves less often than adolescents without ASD (please see Figure 5). Although determining that adolescents with ASD do not visually attend to people for the same amount of time or in the same manner as adolescents without ASD is not a novel finding (Dalton et al., 2005; De Wit et al., 2008; Klin et al., 2002; Tsang, 2018), there is a paucity of data regarding how adolescents with ASD process images of themselves. Results indicated that although the adolescents with ASD appeared reluctant to visually engage with images of people, even themselves, it did not influence their ability to interpret and process associated details. For example, content questions regarding shoe color, jacket color, and sunglasses color were answered accurately and were not related to visual attention metrics. Therefore, adolescents with ASD are able to conceptualize and recall details about people in spite of not engaging in sustained visual attention.

Environment

There were three environmental cues (e.g. weather, a waiting room chair, an office room chair) targeted during each personalized video social story. In contrast to the object and people cues chosen, the environmental stimuli selected as visual cues were deemed extraneous to communicating relevant and salient content. In other words, the environmental visual cue types were included as a way to determine the frequency and duration of visual attention by participants to portions of the social story that were not important. Results indicated that adolescents with and without ASD did not demonstrate significant differences in the amount of time scanning environmental cues or the number of times they focused on environmental cues. However, adolescents with ASD visually fixated on extraneous environmental cues for a longer period of time than on images of people depicted in the social story, including themselves. As such, adolescents with ASD appeared to expend more cognitive

Figure 5. (a) Heat maps looking at facial areas of people without ASD and (b) heat maps looking at facial areas of people with ASD



(a)



(b)

effort on extraneous environmental cues than on relevant person cues within the video social story than adolescents without ASD. Nevertheless, this expenditure of effort did not impact the ability of the adolescents with ASD to identify and remember important elements of the personalized visual social story.

Overall, the results indicated several important points for consideration. During personalized video social stories, adolescents with and without ASD demonstrate differences in visual attention to story content. It appears the specific type of visual stimuli provided impacts the frequency and duration with which adolescents with and without ASD focus on important visual information in a personalized video social story. Specifically, when visual content in personalized social stories are images of people or are person centered, significant differences can be seen in visual attention patterns for adolescents with ASD. Finally, results indicate that, while there are differences in eye gaze patterns during personalized video social stories for adolescents with and without ASD, eye gaze patterns do not relate to the ability to comprehend and maintain novel information presented in the story.

Implications for Practice

The current findings from this study have potential implications for educators and practitioners working with adolescents with ASD. In an effort to increase skills related to predictability and/or theory of mind, clinicians may choose to generate representative event schemas using personalized

video social stories. Adolescents with ASD are able to comprehend and remember details from the real-life event depicted in the personalized social story. Therefore, personalized video social stories are a potentially viable accommodation or intervention for facilitating insight and understanding of event sequences and outcomes. When designing personalized video social stories for adolescents with ASD, attention should be given to the selection of images and visual content. Adolescents with ASD will visually attend more often and longer to images of relevant and important objects encountered during the real-life experience. Even though adolescents with ASD have reduced visual attention to images of themselves and other people, it continues to be important to include and format those images by aligning them with highly preferred objects. Educators and practitioners might consider the placement of objects near images representing persons and people in order to possibly facilitate increased visual scanning to those less preferred visual cues. Additionally, objects that are systematically introduced, relevant, and predictable across scenes will further support the adolescents' ability to generate sustained and regulated eye gaze throughout the video story.

Directions for Future Research

The results of this exploratory study have both research and clinical implications for professionals working with adolescents with ASD. With regard to research implications, the utilization of eye tracking technology during the presentation of dynamic (as opposed to static) stimuli containing personalized images of the participants, that were coded for visual attributes, lends to the novel and innovative nature of this study which should be replicated in future studies. Second, based on the results of this research, it can be posited that eye tracking technology is a feasible method of identifying group differences related to visual attention in adolescents with and without ASD while viewing personalized video social stories. Third, future studies should expand this research to use eye tracking technology to target empirical questions related to determining how adolescents with ASD process images of themselves. Finally, future investigations should examine the ability of adolescents with ASD to process images of self and others in personalized video social stories and whether effective processing impacts the development of skills related to predictability and theory of mind.

Limitations

Despite the overall positive results of this research, there are a number of factors which could have impacted the results of this study. First, the small sample size used for this study potentially limits the generalizability of the findings. However, it is important to point out that since the Tobii eye tracker collects eye gaze data at an approximate rate of 16 milliseconds, four participants yielded a large number of data points. Second, although informal visual analysis of heat maps yielded relevant qualitative information, differences in types of persons and eye gaze metrics for the adolescents who participated in this study were not formally analyzed. Third, it is important to note gender differences in both groups. Specifically, all adolescents with ASD who participated in the study were male, while all adolescents without ASD were female. Again, due to a lack of heterogeneity across participants, caution is advised when considering generalization of the results. Finally, although every effort was made to include an equal number of questions pertaining to each visual cue category in the video social story knowledge measure, each category was targeted only four times. A longer and more detailed measure may have identified significant relationships regarding accuracy and visual attention measures in adolescents with and without ASD.

CONCLUSION

The purpose of this study was to explore eye tracking as a measure of visual attention when watching personalized video social stories for adolescents with and without ASD. Despite the evidence supporting social stories as an effective intervention for reducing disruptive behaviors (Graetz et al., 2009; Scattone et al., 2002), there is a paucity of research regarding how adolescents with ASD process

and comprehend visual content, including images of themselves. The results of this study support the use of eye tracking technology as a physiological method of identifying differences in attention for adolescents with and without ASD. Additionally, this study has shown that personalized video social stories may be effective for adolescents with ASD when the organization and representation of visual images is considered.

REFERENCES

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorder* (5th ed.). American Psychiatric Association.
- Bitouk, D., Kumar, N., Dhillon, S., Belhumeur, P., & Nayar, S. K. (2008, August). Face swapping: Automatically replacing faces in photographs. *ACM Transactions on Graphics*, 27(3), 39. doi:10.1145/1360612.1360638
- Boivin, M. J., Weiss, J., Chhaya, R., Seffren, V., Awadu, J., Sikorskii, A., & Giordani, B. (2017). The feasibility of automated eye tracking with the early childhood vigilance test of attention in younger HIV-exposed Ugandan children. *Neuropsychology*, 31(5), 525–534. doi:10.1037/neu0000382
- Bozkurt, S. S., & Vuran, S. (2014). An analysis of the use of Social Stories in teaching social skills to children with Autism Spectrum Disorders. *Educational Sciences: Theory and Practice*, 14(5), 1875–1892.
- Centers for Disease Control and Prevention. (2018). Prevalence of autism spectrum disorders—Autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *MMWR. Surveillance Summaries*, 67(6), 1–23. doi:10.15585/mmwr.ss6706a1
- Chen, C. H., Bobzien, J., Bruhn, A., Zhang, M., & Hsu, W. W. (2014, October). The first step toward reconstructing visual perception and cognition in Autism. In SMC (pp. 1171-1174). doi:10.1109/SMC.2014.6974072
- Chen, C. H., Bobzien, J., Giannakos, M., Bruhn, A., Bruggeman, A., Mohrehkesh, S., . . . Chrisochoides, N. P. (2015, October). Familiar Video Stories as a Means for Children with Autism: An Analytics Approach. In *Healthcare Informatics (ICHI), 2015 International Conference on* (pp. 368-373). IEEE.
- Crozier, S., & Tincani, M. (2007). Effects of social stories on prosocial behavior of preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37(9), 1803–1814. doi:10.1007/s10803-006-0315-7
- Cuthbert, M. I., & Melamed, B. G. (1982). A screening device: Children at risk for dental fears and management problems. *ASDC Journal of Dentistry for Children*, 49(6), 432–436.
- Delano, M., & Snell, M. E. (2006). The effects of social stories on the social engagement of children with autism. *Journal of Positive Behavior Interventions*, 8(1), 29–42. doi:10.1177/10983007060080010501
- Giannakos, M., Chorianopoulos, K., & Chrisochoides, N. (2015). Making sense of videoanalytics: Lessons learned from clickstream interactions, attitudes, and learning outcomes in a video-assisted course. *International Review of Research in Open and Distance Learning*, 16(1), 260–283. doi:10.19173/irrodl.v16i1.1976
- Graetz, J. E., Mastropieri, M. A., & Scruggs, T. E. (2009). Decreasing inappropriate behaviors for adolescents with autism spectrum disorders using modified social stories. *Education and Training in Developmental Disabilities*, 91–104.
- Gray, C. (2004). Social stories 10.0: The new defining criteria and guidelines. *Jenison Autism Journal*, 15(4), 2–21.
- Gray, C. A., & Garand, J. D. (1993). Social stories: Improving responses of students with autism with accurate social information. *Focus on Autistic Behavior*, 8(1), 1–10. doi:10.1177/108835769300800101
- Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience and Biobehavioral Reviews*, 42, 279–297. doi:10.1016/j.neubiorev.2014.03.013
- Heuer, S., & Hallowell, B. (2009). Visual attention in a multiple-choice task: Influences of image characteristics. *Aphasiology*, 23(3), 351–363. doi:10.1080/02687030701770474
- Karatekin, C. (2004). Development of attentional allocation in the dual task paradigm. *International Journal of Psychophysiology*, 52(1), 7–21. doi:10.1016/j.ijpsycho.2003.12.002
- Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental Review*, 27(3), 283–348. doi:10.1016/j.dr.2007.06.006
- Karatekin, C., Marcus, D. J., & Couperus, J. W. (2007). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. *Psychophysiology*, 44(1), 128–144. doi:10.1111/j.1469-8986.2006.00477.x

- Keehn, B., Brenner, L. A., Ramos, A. I., Lincoln, A. J., Marshall, S. P., & Müller, R. A. (2009). Brief report: Eye-movement patterns during an embedded figures test in children with ASD. *Journal of Autism and Developmental Disorders*, 39(2), 383–387. doi:10.1007/s10803-008-0608-0
- Kokina, A., & Kern, L. (2010). Social story interventions for students with autism spectrum disorders: A meta-analysis. *Journal of Autism and Developmental Disorders*, 40(7), 812–826. doi:10.1007/s10803-009-0931-0
- Kuoch, H., & Mirenda, P. (2003). Social Story™ interventions for young children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 18(4), 219–227. doi:10.1177/10883576030180040301
- McPartland, J. C., Webb, S. J., Keehn, B., & Dawson, G. (2011). Patterns of visual attention to faces and objects in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 41(2), 148–157. doi:10.1007/s10803-010-1033-8
- Melinder, A., Konijnenberg, C., & Sarfi, M. (2013). Deviant smooth pursuit in preschool children exposed prenatally to methadone or buprenorphine and tobacco affects integrative visuomotor capabilities. *Addiction (Abingdon, England)*, 108(12), 2175–2182. doi:10.1111/add.12267
- Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, 31(5), 553–566. doi:10.1068/p3339
- Mondloch, C. J., Robbins, R., & Maurer, D. (2010). Discrimination of facial features by adults, 10-year-olds, and cataract-reversal patients. *Perception*, 39(2), 184–194. doi:10.1068/p6153
- Qi, C. H., Barton, E. E., Collier, M., Lin, Y. L., & Montoya, C. (2018). A systematic review of effects of social stories interventions for individuals with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities*, 33(1), 25–34. doi:10.1177/1088357615613516
- Renzi, C., Schiavi, S., & Carbon, C.-C. (2013). Featural and configural processing of faces are dissociated in the dorsolateral prefrontal cortex: A TMS study. *NeuroImage*, 74, 45–51. doi:10.1016/j.neuroimage.2013.02.015
- Reynhout, G., & Carter, M. (2006). Social Stories™ for children with disabilities. *Journal of Autism and Developmental Disorders*, 36(4), 445–469. doi:10.1007/s10803-006-0086-1
- Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: Evidence from eye tracking. *Developmental Science*, 12(4), 549–556. doi:10.1111/j.1467-7687.2009.00795.x
- Sansosti, F. J., & Powell-Smith, K. A. (2006). Using social stories to improve the social behavior of children with Asperger syndrome. *Journal of Positive Behavior Interventions*, 8(1), 43–57. doi:10.1177/1098300706080010601
- Scattone, D. (2008). Enhancing the conversation skills of a boy with Asperger's Disorder through social stories and video modeling. *Journal of Autism and Other Developmental Disorders*, 38(2), 395–400. doi:10.1007/s10803-007-0392-2
- Scattone, D., Wilczynski, S. M., Edwards, R. P., & Rabian, B. (2002). Decreasing disruptive behaviors of children with autism using social stories. *Journal of Autism and Developmental Disorders*, 32(6), 535–543. doi:10.1023/A:1021250813367
- Serrano, V. J., Owens, J. S., & Hallowell, B. (2018). Where children with ADHD direct visual attention during emotion knowledge tasks: Relationships to accuracy, response time, and ADHD symptoms. *Journal of Attention Disorders*, 22(8), 752–763. doi:10.1177/1087054715593632
- Tsang, V. (2018). Eye-tracking study on facial emotion recognition tasks in individuals with high-functioning autism spectrum disorders. *Autism*, 22(2), 161–170. doi:10.1177/1362361316667830
- Venker, C. E., & Kover, S. T. (2015). An open conversation on using eye-gaze methods in studies of neurodevelopmental disorders. *Journal of Speech, Language, and Hearing Research: JSLHR*, 58(6), 1719–1732. doi:10.1044/2015_JSLHR-L-14-0304

Anne M. P. Michalek (PhD) is a certified speech-language pathologist who has a PhD in Special Education. She is an Assistant Professor in the Communication Disorders and Special Education Department at Old Dominion University.

Jonna Bobzien (PhD) is an Associate Professor and Graduate Program Director at Old Dominion University.

Victor A. Lugo is a speech-language pathology graduate student at Old Dominion University. Prior to attending Old Dominion University, he received his BA in English with a minor in Linguistics from the University of Mary Washington (2012) and served as an Operations Intelligence Analyst in the U.S. Air Force.

Chung Hao Chen (PhD) is current an associate professor in Electrical and Computer Engineering at Old Dominion University.

Ann Bruhn, BSDH, MS, is an instructor at the School of Dental Hygiene, Old Dominion University. Bruhn teaches Radiology and clinical care and has four years experience in clinical practice.

Michail Giannakos (PhD) works at Old Dominion University in the Department of Computer and Information Science.

Anne Chrisochoides (PhD) is a Professor in Computer Science at Old Dominion University.