

Accessible Texts for Autism: An Eye-Tracking Study

Victoria Yaneva
University of Wolverhampton
Wolverhampton, UK
V.Yaneva@wlv.ac.uk

Irina Temnikova
Qatar Computing Research
Institute
Doha, Qatar
itemnikova@qf.org.qa

Ruslan Mitkov
University of Wolverhampton
Wolverhampton, UK
R.Mitkov@wlv.ac.uk

ABSTRACT

People with Autism Spectrum Disorder (ASD) are known to experience difficulties in reading comprehension, as well as to have unusual attention patterns, which makes the development of user-centred tools for this population a challenging task. This paper presents the first study to use eye-tracking technology with ASD participants in order to evaluate text documents. Its aim is two-fold. First, it evaluates the use of images in texts and provides evidence of a significant difference in the attention patterns of participants with and without autism. Sets of two types of images, photographs and symbols, are compared to establish which ones are more useful to include in simple documents. Second, the study evaluates human-produced easy-read documents, as a gold standard for accessible documents, on 20 adults with autism. The results provide an understanding of the perceived level of difficulty of easy-read documents according to this population, as well as the preferences of autistic individuals in text presentation. The results are synthesized as set of guidelines for creating accessible text for autism.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Screen design*; K.4.2 [Computers and Society]: Social Issues—*Assistive technologies for persons with disabilities*

Keywords

Autism, Readability, Reading Comprehension, Text Simplification, Easy-Read, Eye Tracking

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterised by impairment in communication and social interaction and stereotyped repetitive behaviour [2]. It is a condition of high prevalence, affecting about 1 in 100 people in the UK [8], and it is considered that for every three known cases, there are two undiagnosed individuals

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org
ASSETS'15, October 26–28, 2015, Lisbon, Portugal.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-3400-6/15/10 ...\$15.00.

DOI: <http://dx.doi.org/10.1145/2700648.2809852>

who might need a diagnosis at some point in their lives [4]. Broadening the diagnostic criteria and raising awareness of these criteria are expected to result in an even higher number of autism diagnoses, especially among women and girls. Impairment in communication, both in terms of language comprehension and social interaction, is a central characteristic of autism. People who have ASD experience various problems in reading comprehension, such as inability to understand context, process long and complex sentences and comprehend figurative language and abstract words [19, 21, 33].

Attention also develops differently in those with autism, with a record of atypical attention patterns dating back to as early as the first mention of this condition by Leo Kanner in 1943 [26]. For instance, the autistic individual may rely on only one sensory modality, while several are relevant to a task — a phenomenon known as stimulus overselectivity [29]. In the context of a reading task, some autistic people might focus exclusively on small fragments of local information with less account for global, contextual and semantic information [22]. These reading difficulties have been pointed out as a reason for the lower educational achievements of students with ASD [9], resulting in future problems with employment; only 6% of the adults with autism in the UK are employed full-time [31].

To address this growing issue there are a number of tools designed to facilitate written communication in people with autism (Section 2.1). These tools are developed for autistic people with a wide range of abilities, starting with severely impaired individuals who could use images in Picture Exchange Communication Systems (PECS) to construct sentences. At the high end, there is text simplification software such as OpenBook [3] for relatively high-functioning individuals, who only need assistance with certain linguistic constructions, such as splitting difficult sentences, accessing a word definition, or having a concept illustrated by an image (Section 2.1). An alternative to computer-oriented solutions are human-produced easy-read documents, following a pre-defined set of guidelines for writing in Plain English [17, 40]. Regardless of the media used to make the text accessible, the output of most computer-based systems and Plain English guidelines has the following common characteristics: it is a document containing short paragraphs of simple text, accompanied by an image related to the meaning of the paragraph or to some of the words in it. However, up to the present point there have not been any studies investigating the influence of text and image combinations on the atypical attention in autistic readers, and whether

these readers make use of the image as a comprehension cue. Even though autistic people are known to have difficulties inferring meaning from symbols and drawings as opposed to photographs (Section 2.3), currently both types of images are widely present in easy-read documents (Section 3). Furthermore, there is no information on whether or not autistic adults with no developmental delay prefer to have images inserted in the texts they read and whether they perceive the level of difficulty of Plain English texts as suitable for their comprehension skills. This paper presents the first study to use eye-tracking technology with ASD participants in order to evaluate text documents. The main contributions of the study are as follows:

- The study investigates the differences between autistic and non-autistic individuals in the proportion of time they spend looking at images and text paragraphs, with a view to improving the development of user-centred document layouts and interfaces.

- It tests the existing assumption (Section 2.3) that autistic people are more efficient in decoding images with high resemblance to their referent in reality (photographs), rather than images with low resemblance to the referent (symbols) on a group of adults with autism with no developmental delay. Confirmation that such difference exists would cast light on the most suitable set of images to be used in the development of accessible reading documents for adults with autism, as well as in other software designed for this population.

- It tests the perceived level of difficulty of Plain English texts on adults with autism, in order to establish whether they are perceived as too difficult or too easy, thus potentially leading to the participants losing interest in the reading material and having their concentration reduced.

- Finally, the study contributes to a greater awareness of the participants' own preferences on text presentation with regard to the use of images.

The next section presents related work on tools for aiding language comprehension in individuals with ASD (Section 2.1) and image processing in autism (Sections 2.2 and 2.3). Section 3 focuses on the common characteristics of human-produced easy-read documents on the web. Sections 4 and 5 describe the hypotheses and experimental methodology, the results of which are given and discussed in Sections 6 and 7. The guidelines for improving accessibility of text documents for readers with autism are presented in Section 8 and conclusions are summarised in Section 9.

2. RELATED WORK

A number of studies report that software and technologies are well-received among autistic individuals due to reasons such as the autistic need for sameness and structure, the ability of the tools to continuously repeat given actions and instructions and, last but not least, the fact that they lack the complexity of social situations [35, 7].

2.1 Language Assistance Tools for Autism

Recognising the need of autistic individuals for assistance with language tasks, there are a number of software tools

developed to assist both children and adults with autism from a broad range of linguistic and intellectual abilities. Probably the most intuitive example of such software are the various types of PECS, where the person produces a sentence by combining different images or images with words on a tablet screen or a personal digital assistant (PDA) [10]. Another tool which helps with language production is the *VAST-Autism*¹ app, which combines videos with written words and auditory cues to help autistic and apraxic individuals acquire certain words, phrases or sentences, similar to the way *Mind Reading*² uses a library of 412 basic human emotions to teach autistic individuals to recognise facial expressions. Another iPad application which helps the expressive vocabulary of children, autistic people and people with other special needs is *Stories About Me*³, allowing them to combine photos with text and voice recordings, in order to produce their personal story. In terms of facilitating receptive language skills, the tool *OpenBook*⁴ [3] provides automatic conversion of text documents by allowing the users and their carers to perform operations such as sentence splitting, inserting word definitions, explaining the meaning of idioms, replacing a word with a synonym, customising the text and page layout, etc. A text summarisation option is also available to provide the users with short versions of the document, and image insertion is used to illustrate unknown words or to reinforce the retention of recently acquired ones. The tool is available for English, Spanish and Bulgarian and has been evaluated on both children and adults.

A common feature in all tools mentioned above is the use of images to facilitate the cognitive processing of information. Nevertheless, information on the use of images in software interfaces is scarce. Some user requirement surveys stress that no bright colours or background images should be used [34]. A survey with 120 autistic respondents and their families concludes that sensory integration and attention issues should be addressed by allowing users to set colours or sounds [35], but it did not investigate preferences on the use of images or visual cues in software. The next section discusses current research on the way people with ASD process images with high and low resemblances to their intended referents.

2.2 Photographs and Symbols

The Oxford Dictionary of English defines the word symbol as “a mark or character used as a conventional representation of an object, function, or process”⁵. The cognitive processing of symbols and photographs, as, respectively, weak and strong representations of their referred objects, requires two different levels of symbolic understanding. In the process of childhood development, children first play with real objects and learn to associate them with the activity they are used for. When they learn to match the real object with a photograph of it, they demonstrate a higher level of symbolic understanding, the next step of which is learning to

¹<https://itunes.apple.com/us/app/vast-autism-1-core/id426041133?mt=8>

²<http://www.jkp.com/mindreading>

³<https://itunes.apple.com/us/app/stories-about-me/id531603747?mt=8>

⁴<http://www.openbooktool.net/>

⁵Oxford English Dictionary [Online] Available at: <http://dictionary.oed.com>



Figure 1: Examples of a symbol and text pair (above) [32] and a photograph and text pair (below) [11]

match the object to a drawing or a symbol and which culminates in the acquisition of a word to denote this whole set of entities [12, 6]. In the case of easy-read documents, both images with a strong resemblance to their intended referent (photographs) and images with a low resemblance (drawings and symbols) have been widely used (Figure 1). Without entering the philosophical debate on the differences between symbols and signs, the current paper uses the term “symbol” in order to refer to images in adapted documents, different from photographs of real people or objects, for example.

2.3 Symbolic Understanding in Autism

Children with autism are considered to have greater difficulties decoding vague representations, compared to typically developing individuals, due to their impaired ability to generalise, grasp context, or reason about the intention of the author [23, 1]. This evidence suggests that the type of images used would have a greater impact on the perception of the autistic users than on that of the neurotypical (non-autistic) ones. In line with this assumption, some developers of language assistance software for people with autism highlight challenges such as “the issue of identification of the most appropriate set of pictures for this system” [37] and hypothesise that:

In [the] case of a child with autism, due to their difficulty with abstraction and generalization, the pictures need to have a strong resemblance to their referents. The more relevant these pictures are to the child’s culture and environment, the easier it is for them to use the system. [37] (p.35)

However, until now, this assumption has not been evaluated and both images with strong and weak resemblances to their referents (photographs and symbols) have been widely used. Furthermore, there is no evidence whatsoever on whether this difference in perception continues in adulthood and whether adult users with autism take longer to process images with a low resemblance to their referent.

One way to test this is to use overt visual attention as a proxy to the cognitive load each set of images imposes on the participants. The term overt attention refers to attention which is observable (e.g. where and when the eyes are changing their position on the screen) as opposed to covert attention, which is not directly observable (e.g. the actual cognitive processing of what the eyes are (or are not) looking at). Overt attention has been used for decades to investigate the cognitive load imposed on participants in different tasks such as reading, watching videos, using website menus, etc.

[39, 13, 18]. In this paper we use overt visual attention in order to establish which set of images, photographs or symbols, imposes a greater cognitive load on the participants, as well as to identify any differences in the attention patterns between the two groups in a task of reading easy-read documents.

3. EASY-READ TEXTS ON THE WEB

Easy-read documents are documents specifically produced for people with comprehension difficulties by following a set of guidelines for writing in Plain English [17, 40]. A document is considered written in Plain English when its readability level is higher than 65 according to the Flesch Reading Ease Scale [15] or at a corresponding level according to other readability formulae. In order to find out the common characteristics of easy-read documents on the web, available to people with special needs, a pool of 100 randomly selected easy-read documents was created, consisting of an overall 71,627 words and 11,522 sentences.

Table 1 summarises some of the characteristics of the sample, where Flesch Reading Ease: 0=very difficult and 100 = very easy [15]; Flesch-Kincaid Grade Level (**FKGL**): 0 = very easy and 12=very difficult [27], **SL** stands for sentence length in words, **WL** for word length in syllables, **AA** for age of acquisition of the content words and **F** for familiarity. All measures have been obtained using the Coh-Metrix 3.0 software [30].

Table 1: Characteristics of 100 easy read documents

	SL	WL	AA	F	Flesch	FKGL
Mean	6.6	1.4	314	581	78	4
Min	2.8	1.2	236	566	43	0.3
Max	15	1.9	410	596	100	10
SD	2.6	0.1	35	6.3	12	2

The documents included in the pool were obtained from various charity organisation websites (n=38), government departments (n=26) and healthcare services (n=32), as well as demos of adapted books (n=3) from the US and the UK. All documents were written in English.

While still providing valuable information about the level of difficulty of a text, readability formulae are not entirely suitable for evaluation of easy read materials due to the great number of stand-alone words, email addresses, links to websites, etc. Furthermore, the choice of appropriate pictures, the layout of the text and the appropriateness of the material for the intended audience cannot be measured by the currently existing formulae. In the present study, human-produced easy-read documents have been chosen as a gold standard for accessible documents. This was partly done in order to allow generalisation of the results across a broader range of tools and human-produced documents instead of evaluating the performance of only one of them, and partly because people with special needs have much wider access to such documents compared to their access to specialised software.

4. HYPOTHESES

The participants in the study were 20 people diagnosed with autism and 20 non-autistic control subjects, who all read 9 texts, while having their eye movements recorded by an eye

tracker.

The study investigates whether there are any between-group differences in the proportion of time each group spends looking at the image in 39 text and image pairs (see Figure 1). The study also investigates which type of images, photographs or symbols, elicit longer fixation times and thus impose a heavier cognitive load on the participants. We also want to find out how the level of difficulty of easy-read documents is perceived by adults on the spectrum. Comparing the perceived level of difficulty of the 9 texts presented would also give information on whether texts considered to be written in Plain English but have different readability levels would evoke different responses from the participants. Finally, we investigate what the text presentation preferences of the two groups are, by including a survey question at the end of the experiment. These research questions are summarised in the following 4 hypotheses:

H.1: There is no difference between groups on the proportion of time spent looking at the image for each text and image pair.

H.2: For each group, there is no effect on the time spent looking at photographs and symbols.

H.3: There is no difference between groups on the perceived level of difficulty of the presented documents.

H.4: There is no difference between groups regarding the text presentation preferences.

The design and procedure of the experiment testing these hypotheses are presented in Section 5.

5. EXPERIMENTAL METHODOLOGY

5.1 Design

The study implemented both between-group and within-group comparison design, where the independent variable was the use of images in texts and had three levels: texts with photographs (20 photographs in total), texts with symbols (19 symbols in total) and plain texts (with no images). After reading each text and having their reading fixations recorded by an eye tracking device, the participants were asked one literal multiple choice question about its meaning, in order to ensure that they were reading for meaning as opposed to just skimming through the text. The questions testing the comprehension of the participants were chosen to be literal due to the simplicity of the easy-read texts, where, by default, no strong inferential or reorganisational skills are needed in order to comprehend their meaning. As an example, a text about eating habits, where various types of foods were discussed, would be followed by a literal multiple choice question with only two possible answers:

High fibre foods include: a) Meat and milk or b) Bread and beans?

Knowing that they would need to answer a question after the text is removed from sight, all participants read the documents carefully (as evidenced by the gaze pattern videos produced by the eye tracker) and were all able to answer 100% of the questions correctly. The answers to these questions are used as a control variable only and are not included

in the analysis of this study. After reading each text and answering the multiple choice questions, participants would rate their subjective perception of the difficulty of the text on a Likert scale from 1 to 5, where 1 stood for “very easy” and 5 stood for “very difficult”. Finally, all participants answered a question about their reading preference, where they could choose between reading texts with a) photographs, b) symbols, c) plain text (no images) or d) “It makes no difference to me”. Participants were allowed to choose more than one or none of the answers and were encouraged to elaborate on their choice if they wanted to. Based on the above design, we considered 6 metrics in total. Images and text paragraphs were defined as areas of interest (AOIs) and a number of gaze-based metrics were obtained based on how many times and for how long participants looked at these areas. The metrics used in this study are:

Average Time Viewed (ATV): The average time an AOI was viewed by all participants measured in seconds. This is an average from the total dwell time, including the durations of all fixations and all revisits.

Average Number of Fixations (AF): The average number of gaze fixations from all participants in a given AOI.

Average Number of Revisits (AR): The average number of go-back gaze fixations from all participants in a given AOI. Go-back gaze fixations are fixations in the span of a given AOI elicited after the gaze path has left the AOI and has then returned to revisit it. Revisits are a valuable source of information for heavy cognitive load and the way information from different parts of the screen is integrated.

Reading time score: This measure was developed by estimating the mean reading time per text in each group and then dividing the result by the number of words in the text. This was done in order to control for the differences in length between the 9 texts. Reading time has been used as an indicator of reading difficulty, with examples of texts similar in length but differing in the time they require for reading based on their complexity level [14].

Perceived level of difficulty: This measure was obtained through 1-5 Likert items after each text where the participants would rate the level of difficulty of the text as they have perceived it. This measure was chosen instead of reading comprehension questions as it more accurately reflects the subjective impressions of the participants on text difficulty and is thus more useful for evaluating their attitudes towards the difficulty of Plain English texts.

Text presentation preference: Information was gathered through the following survey question: “In your everyday life, do you prefer reading texts with: a) photographs, illustrating the main ideas b) symbols, illustrating the main ideas, c) plain texts without any images or d) “It makes no difference to me”.

5.2 Participants

20 adults (7 female, 13 male) with a confirmed diagnosis of autism (n=10 Autism Spectrum Disorder, n=9 Asperger’s syndrome and n=1 semantic-pragmatic disorder) were recruited through 4 local charity organisations. Participants

in the control group were 20 non-autistic adults (11 female and 9 male). None of the 40 participants had comorbid conditions affecting reading (e.g. dyslexia, learning difficulties, aphasia etc.), but some participants were diagnosed with comorbid depression (n=4, ASD group; n=1, control group) and anxiety (n=6, ASD group). Mean age (m) for the ASD group was m=30.75, with standard deviation SD=8.23, while for the control group it was m=30.81, SD=4.8. Years spent in education, as a factor influencing reading skills, for the ASD group were m=15.31, SD=2.9, and for the control group, m=17.25, SD=2.15. None of the participants in the two groups were diagnosed with having a learning disability or a developmental delay, so no matching for mental age was required [25]. All participants were native speakers of English and had normal or corrected vision. Results from 3 participants from the ASD group were discarded due to poor calibration or data loss (too many head movements during reading), resulting in dramatic inaccuracies in more than 70% of the data collected from them. Hence, the results analysed were obtained from 17 ASD and 20 control participants.

5.3 Materials

A challenge to the design of this study was the selection of an appropriate number of texts to give sufficient statistical power to the experiment without imposing a heavy work load on the participants with ASD. The inclusion of a large number of test materials was not feasible, due to the fact that people with autism are known to experience difficulties with concentration and attention [9, 38] and sensory issues [16, 28], in addition to the longer time they take to comprehend instruction, to calibrate the eye tracker [36], and to get accustomed to new environment, such as the room where the experiment takes place [5]. The materials used in the study were easy-read documents. In order to ensure that the texts included in this study were representative of the easy-read information available to a person with special needs, they were selected from the pool of 100 easy-read documents discussed in Section 3. Out of the 100-document pool, a sample of 7 texts comprising of 39 image and text snippets was carefully chosen for the experiment, based on the following criteria: **Topic** (all documents included did not require any prior knowledge nor did they discuss sensitive topics), **Source** (the selected documents covered all sources listed above, such as charity organisations, government and healthcare departments), **Readability level** (documents (or parts of documents) were included so as to cover a diverse number of readability levels), and **Images** (both photographs (n=20) and symbols (n=19), were included, each of them accompanied by paragraphs of text as opposed to one- or two-word descriptions). As the easy-read documents contain images by default, the 2 texts in the condition “plain text without images” were selected from the WeeBit readability corpus [41]. They were also written according to Plain English requirements and their readability scores were medium compared to the 7 easy-read documents selected. Thus, the study included 9 texts overall, the information of which is summarised in Table 2. Readability scores have been obtained using the Coh-Metrix 3.0 software [30].

5.4 Apparatus

The device used for recording the gaze of the participants during task performance was a Gazepoint GP3 video-based

Table 2: Characteristics of the texts included in the experiment

	Type	Words	Images	FKGL	Flesch
T1	Photos	77	4	8.16	60.11
T2	Photos	96	5	6.73	67.33
T3	Symbols	74	6	2.71	92.54
T4	Photos	178	8	5.52	75.33
T5	Symbols	77	6	5.79	70.67
T6	Symbols	121	6	1.75	95.00
T7	Photos	58	4	6.63	68.16
T8	None	178	0	4.67	80.22
T9	None	163	0	4.93	79.548

eye tracker [20] (60Hz sampling rate), with a 19” LCD monitor. No equipment was attached to the participants. The eye tracker was calibrated individually for each participant using a 9-point calibration procedure. The use of a chin rest is not recommended due to sensory issues common within autism [38], which is why the distance between each participant and the eye tracker was controlled by using a fixed chair only, and was roughly 85 cm.

5.5 Procedure

All participants performed the experiment in a quiet room with only the researcher (first author) present. Some of them (n=2) requested to have the lights diminished due to sensory issues, which was done in order to ensure their comfort during the experiment. First, each participant was given a verbal instruction on the sequential order in which the experiment was going to proceed and on the functionality of the eye tracker. Each participant was given the opportunity to ask questions and to request a break at any point if they felt tired. Recalibration was performed if the participants needed to get up during their breaks. Demographic information was collected after the instruction was given. After that the instruction was reinforced a second time and the calibration procedure started. Each of the documents was presented on screen in a randomised order and participants could take as long as they needed to read it. After each document the participant would be asked a comprehension question verbally, without having the opportunity to look at the text while answering. At the end of the experiment the survey question was asked and participants were debriefed.

6. RESULTS

The software used for the analysis of the fixation points from the eye tracker and their grouping into specific areas of interest (AOIs) such as images or text paragraphs, was the Gaze-point analysis system, specifically developed for the GP3 Gazepoint eye trackers [20]. Statistical data was analysed using IBM SPSS Statistics software, Version 20 [24].

6.1 Attention to Images

H.1: *There is no difference between groups on the proportion of time spent looking at the image for each text and image pair.*

First, we compared the overall time participants from both groups spent looking at the 39 images. A Shapiro-Wilk test showed that all the gaze-based measures datasets, namely Average Time Viewed (ATV), Average Fixations (AF) and

Average Revisits (AR), are non-normally distributed. Hence, the study used a Mann-Whitney U test to assess the null hypothesis according to all three gaze-based measures. The test significantly rejected this hypothesis confirming a difference between the groups, where the participants with autism were shown to spend significantly longer time not only looking at the images, (ATV: $U=338.5$, $N_1=17$, $N_2=20$, $p=0.000$, two-tailed; AF: $U=290.000$, $N_1=17$, $N_2=20$, $p=0.000$, two-tailed) but also revisiting them (AR: $U=331.000$, $N=17$, $N_2=20$, $p=0.000$, two-tailed). However, a significant difference between the absolute times participants spend looking at the images may result from longer overall reading times in the ASD group. To investigate this further, for each group we added up the average viewing times (ATV) for each image together with the ATV of its corresponding paragraph, so that we had AOIs containing the ATV for 39 text and image pairs. Then the ATV of each image was computed as a percentage of the time for each pair (taken as 100%) in the following way: $ATV \text{ per image } (\%) = ATV \text{ per image and text pair} - ATV \text{ per text paragraph}$. A Mann-Whitney U test indicated that there is a statistically significant difference between the two proportions ($U=461.00$, $N=17$, $N_2=20$, $p=0.003$, two-tailed), and thus **H.1** was rejected, with the ASD group spending a greater proportion of time on images compared to the control group, which is evidence of an atypical attention pattern in this population. The proportion of time the ASD group spent looking at the images totalled 20.32%, compared with 13.42% for the control group, leaving the ASD group with 79.68% of their time spent on reading the text and 86.58% for the control group.

6.2 Photographs vs. Symbols

H.2: *For each group, there is no effect on the time spent looking at photographs and symbols.*

A Shapiro-Wilk test showed that the data is non-normally distributed for Average Time Viewed (ATV) (Symbols: $p=0.001$, Photographs: $p=0.011$) and Average Revisits (AR) (Symbols: $p=0.000$, Photographs: $p=0.163$) in the control group, while the Average Fixations (AF) dataset for both control and ASD group and the datasets ATV and AR for the ASD group were normally distributed (Control group: Symbols $p=0.001$, Photographs $p=0.011$; ASD group: Symbols $p=0.091$, Photographs $p=0.332$). Hence, a paired-samples t-test was used to compare the data in the “symbol” and “photograph” classes for the ASD group for all three measures and for the AF dataset from the control group. A Wilcoxon Matched Pairs test was in turn used to compare the non-normally distributed ATV and AR datasets for the control group. First a test for outliers was performed, which showed that there were no outlier values in the datasets. The paired-samples t-test showed that in the ASD group there was no significant difference between the time spent viewing images according to the ATV measure ($t=-1.389$, $df=18$, $p=0.182$, two-tailed), AF ($t=-1.339$, $df=18$, $p=0.197$, two-tailed) or AR ($t=0.378$, $df=17$, $p=0.710$, two-tailed). Similarly, the results from the Wilcoxon Matched Pairs test revealed that there is no significant difference between the times participants in the control group spent looking at symbols or photographs for the ATV and AR measures (ATV measure: $z=-0.765$, $N-Ties=19$, $p=0.444$, one-tailed; AR measure: $z=-0.763$, $N-Ties=17$, $p=0.445$, one-tailed), and a paired-samples t-test confirmed the same for the AF measure ($t=-0.298$, $df=18$, $p=0.769$).

The results failed to reject the H.2 hypothesis that there is no significant difference between the times participants from both groups look at photographs and symbols. The results indicate that photographs and symbols impose similar cognitive loads on the participants from both groups and thus both sets are equally suitable for use in easy-read documents for adults with ASD.

6.3 Level of Difficulty

H.3: *There is no difference between groups on the perceived level of difficulty of the presented documents.*

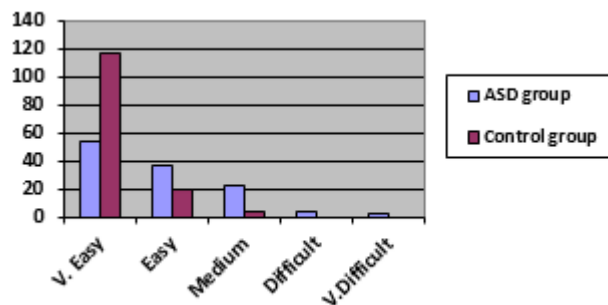


Figure 2: Frequency of text rankings for each category

A Shapiro-Wilk test showed that for the ASD group the data was not normally distributed for all texts, with the exception of Text 8, and for the control group the data was not normally distributed for all texts. Hence, to study the occurrence of any significant differences between the perceived level of difficulty in the 9 texts (Figure 2), we used a Friedman’s non-parametric test for repeated measures, which showed no statistically significant difference between the perceived level of difficulty in both groups for all 9 texts (ASD group: $\chi^2(8) = 9.679$, $p=0.139$, control group: $\chi^2(8) = 10.145$, $p=0.119$), indicating that documents at readability levels between 61 and 95 Flesch Reading Ease score are considered as the same class of difficulty by the ASD group.

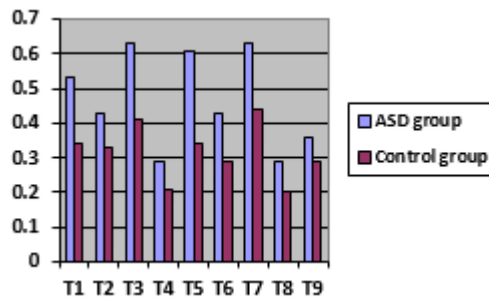


Figure 3: Differences in reading time scores between the autistic and non-autistic participants

Nevertheless, there were expected between-group differences in the reading time for each document, showing that despite the lack of developmental delay, the ASD group did struggle more with reading the 9 texts (Figure 3). Furthermore, the ASD group showed more diverse answers to the perceived

level of difficulty of the texts with answers ranging between very easy (n=54), easy (n=37) and medium (n=23) to even reaching difficult (n=4) and very difficult (n=2). The control group, on the other hand, had a predominant rate of very easy (n=117) and easy (n=20) and none of the participants ranked any text as difficult or very difficult (Figure 2).

6.4 Text Presentation Preferences

H.4: *There is no difference between groups regarding the text presentation preferences.*

There was a strong preference for the inclusion of images among the ASD group (58.81%), of which 23.5% preferred reading texts with photographs and 35.3% preferred reading texts with symbols. The control group did not declare such a strong preference to images with 60% of the participants stating that it makes no difference to them and 30% voting for the inclusion of images but were undecided as to whether they preferred photographs (15%) or symbols (15%) (Figure 4).

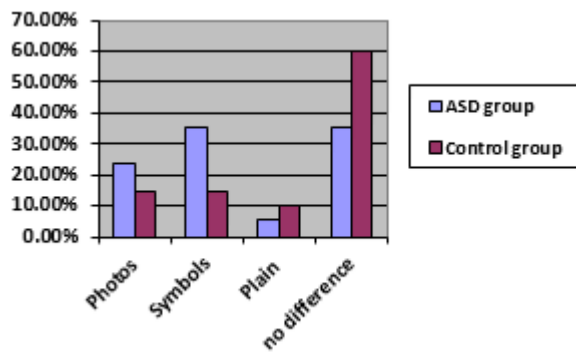


Figure 4: Preferences of the two groups on text presentation

7. DISCUSSION

The results indicate that there is a significant difference between autistic and non-autistic individuals in the proportion of time they spend looking at images and text paragraphs, with autistic participants relying on the image more than the non-autistic ones. This result may have an implication on the design of user-centred software in two ways.

First, it is evidence that visual cues take up a significant part of the attention of autistic participants and that they would make useful comprehension cues if they are closely related to the meaning of the text paragraph they accompany.

Second, it confirms that there are differences in attention between neurotypical and autistic people with regards to images and that the improper inclusion of images in text documents or software interfaces may have a negative effect by distracting the autistic users much more than the intuition of the neurotypical developers may suggest.

Both photographs and symbols have been found to elicit similar cognitive load on the participants, which means that in the case of adult autistic users without developmental delay, both sets of images are equally suitable to be included. This finding is not in conflict with previous research, as it is the only study on images so far that has included adult

autistic participants instead of children. It is possible that there would be a significant difference between the two sets if the participants were children, as it may be the case that symbolic understanding in autistic individuals reaches levels equal to those of neurotypicals later in their lives. In this sense, the results of this study with regard to the types of images should not be generalised to children or to autistic individuals with learning difficulties.

The fact that texts written in plain English are perceived as ranging from very easy and easy to difficult and even very difficult by the participants with autism, while non-autistic participants rate them predominantly as very easy, is an indication that these texts are well understood by the autistic participants without being as trivial to them as they have appeared to be for the non-autistic ones. Perceived level of difficulty is not a direct measure of interest but one could hypothesise that texts which are too easy would bore the readers and thus reduce their motivation to read the document. The results suggest that even though our autistic participants were adults without a learning disability, this is not the case with them, and that Plain English texts are indeed a suitable level of difficulty for this population.

One factor which may have influenced these results is that the study did not use deception and all participants knew that it investigated reading in autism. The autistic participants might have suggested they were expected to have some sort of reading deficits and have thus tried to apply a more fine-grained classification of the difficulty of the texts compared to the non-autistic ones. Nevertheless, differences in the interpretation of Likert items are a well-known flaw in all types of studies using this measure, while in the case of this study the results from the Likert scale are in agreement with the longer reading times of the autistic participants, which support the conclusion that they indeed did not find the texts as easy as the control participants did.

The slight majority of the autistic participants preferred to have images included in the documents they read (58.81%), with 23.52% of them preferring photographs and 35.29% preferring symbols, while roughly the same proportion of the non-autistic participants were indifferent as to whether images were included in the text or not (60%). This is another difference between neurotypical and autistic individuals, which should be taken into account when creating accessible texts for this population, even when the target group is adults.

Limitations of this study are the relatively small number of documents assessed and the small number of participants. The first one is imposed by the difficulty experienced by autistic participants in concentrating for long periods of time, as discussed in Section 5, and could only be overcome if the participants come for several shorter sessions. The second limitation is typical of all areas of autism research, which makes the results from these studies difficult to generalise and is the reason for the many inconsistencies in autism study replications. The reason that the samples in ASD research tend to be small is the varying levels of ability of this population and the number of comorbid disorders, (learning difficulties, dyslexia, depression, apraxia) which are so common within autism but in many cases need to be excluded for the purposes of research.

8. IMPROVING TEXT ACCESSIBILITY FOR READERS WITH AUTISM

The results of the study presented above are synthesized in a set of recommendations regarding text accessibility for adult readers with autism. These recommendations aim to improve the development of text documents, Web content or software interfaces targeted to autistic adults, as well as text conversion and language assistance tools, such as the ones described in Section 2. The recommendations based on the results from this study are as follows:

1. Illustrate the main ideas in text paragraphs through the insertion of images relevant to the meaning of the paragraph.
2. If a relevant image is unavailable or the idea of the text is too abstract to be depicted as an image, do not put anything. A non-relevant image has the potential to affect autistic readers' comprehension and reading speed.
3. Do not insert logos, advertisements or any other visual information, which is not directly relevant to the meaning of the text.
4. Insert the image as close as possible to the sentence or groups of sentences it refers to.
5. Photographs and symbols are equally suitable, so data sets from both domains could be utilized. However, refrain from using symbols which are too abstract or their understanding requires substantial prior knowledge in a certain area.
6. Use texts written in plain English. See plain English guidelines for a more detailed information on how to write for people with disabilities [40]. A general rule of thumb is that the text should have a score higher than 65 according to the Flesch-Reading Ease formula [15].
7. Allow the readers to skip through the pages at their own pace, as reading times may be longer compared to the general population.
8. In the case of videos, allow longer times for the users to read the text or captions and to process the visual information.

It is important to note that due to the heterogeneity of autism, these recommendations should be applied with caution and should not be generalised to children with autism or adults at the lower ends of the autism spectrum where signs of a learning disability are evident.

9. CONCLUSIONS AND FUTURE WORK

The study provided evidence for the differences in attention patterns between participants with and without autism and suggested that both photographs and symbols are equally useful for including in easy-read documents and software interfaces. Documents written in Plain English are understood by all participants with autism but are not all ranked as 'very easy'. This suggests that this level of difficulty was suitable for the autistic participants to understand without

being too trivial for them, even though they were adults and had no learning difficulties. Finally, the study indicates that the majority of people with autism prefer to read texts with images, unlike neurotypicals, for whom it mostly makes no difference whether there are images in the text or not. An interesting direction for future work is to analyse the types of visual cues which aid comprehension for autistic adults, as well as to analyse their gaze fixations for particular constructions in the texts.

10. ACKNOWLEDGEMENTS

The authors would like to thank all participants who took part in the study, as well as the charity organisations, which facilitated their recruitment. Special thanks to the three anonymous reviewers for their constructive criticism and to Dr. Luz Rello for her valuable help and advice.

11. REFERENCES

- [1] M. L. Allen. Brief report: decoding representations: how children with autism understand drawings. *Journal of autism and developmental disorders*, 39(3):539–43, Mar. 2009.
- [2] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (5th ed.), 2013.
- [3] E. Barbu, M. T. Mart, and L. A. Ure. Open Book : a tool for helping ASD users' semantic comprehension. In *Proceedings of the 2th Workshop of Natural Language Processing for Improving Textual Accessibility (NLP4ITA)*, number June, pages 11–19, Atlanta, Georgia, 2013. Association for Computational Linguistics.
- [4] Baron-Cohen, S., Scott, F. J., Allison, C. Prevalence of autism-spectrum conditions: UK school-based population study. *British Journal of Psychiatry*, 194:500– 509, 2009.
- [5] Bejerot, S., Eriksson, J. M. and Mörtberg, E. Social anxiety in adult autism spectrum disorder. *Psychiatry research*, 220((1-2)):705–7, 2014.
- [6] E. Bialystok. Symbolic representation across domains in preschool children. *Journal of experimental child psychology*, 76(3):173–89, 2000.
- [7] Bosseler, A and Massaro, D. W. Development and evaluation of computer-animated tutor for vocabulary and language learning in children with autism. *Journal of autism and developmental disorders*, 33(6):553–567, 2003.
- [8] Brugha, T., Cooper, S. A., McManus, S., E. Estimating the Prevalence of Autism Spectrum Conditions in Adults: Extending the 2007 Adult Psychiatric Morbidity Survey. Technical report, NHS, The Health and Social Care Information Centre., London, 2012.
- [9] Brugha, T., McManus, S., Meltzer, H., Smith, J., Scott, F.J., and Purdon, S. Autism spectrum disorders in adults living in households throughout England. Report from the Adult Psychiatric Morbidity Survey 2007. Technical report, The National Health Service Information Centre for Health and Social Care, London, 2007.
- [10] Charlop-Christy, Marjorie H., Carpenter, Michael, Le, Loc, LeBlanc, Linda A. and Kellet, Kristen. Using the picture exchange communication system (pecs) with

- children with autism: assessment of pecc acquisition, speech, social-communicative behavior, and problem behaviour. *Journal of Applied Behavior Analysis*, 3(3):213–231, 2002.
- [11] Cheshire and Wirral Partnership NHS Foundation Trust. Healthy Eating and Being Active. Information on Diet and Exercise. Easy read leaflet, 2008.
- [12] J. S. DeLoache. Symbolic Functioning in Very Young Children: Understanding of Pictures and Models. *Child Development*, 62(4):736–752, 2008.
- [13] V. Demberg and F. Keller. Data from Eye-tracking Corpora as Evidence for Theories of Syntactic Processing Complexity. *Cognition*, 109(2):193–210, 2008.
- [14] W. H. Dubay. The Principles of Readability. 2004.
- [15] R. Flesch. A New Readability Yardstick. *Journal of Applied Psychology*, 32(3):221–233, 1948.
- [16] Freeman, B. J., Ritro, E. R. and Schroth, P.C. Behavior Assessment of the Syndrome of Autism: Behavior Observation System. *Journal of the American Academy of Child Psychiatry*, 23:588–94, 1984.
- [17] Freyhoff, G., Hess, G., Kerr, L., Tronbacke, B. and Van Der Veken, K. Make it Simple. European Guidelines for the Production of Easy-to-Read Information for People with Learning Disability. Technical report, ILSMH European Association., 1998.
- [18] S. Frisson and M. J. Pickering. Obtaining a Figurative Interpretation of a Word : Support for Underspecification. 16:149–171, 2001.
- [19] Frith, Uta and Snowling, M. Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology*, 1:329–342, 1983.
- [20] Gazepoint. GP3 Gazepoint Eye tracker, 2015.
- [21] F. Happe. Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology*, 15:1–12, 1997.
- [22] Happe, F., and Frith, U. The weak coherence account: Detail focused cognitive style in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 36:5–25, 2006.
- [23] C. Hartley and M. L. Allen. Intentions vs. resemblance: understanding pictures in typical development and autism. *Cognition*, 131(1):44–59, Apr. 2014.
- [24] IBM Corp. IBM SPSS Statistics for Windows, Version 20.0, 2011.
- [25] Jarrold, C. and Brock, J. To Match or Not To Match ? Methodological Issues in Autism Related Research. *Journal of autism and developmental disorders*, 34(1):81–86, 2004.
- [26] L. Kanner. Autistic Disturbances of Affective Contact. *Nervous Child*, 2:217–250, 1943.
- [27] Kincaid, J. P., Fishburne, R. P., Rogers, R. L., and Chissom, B. S. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. Technical report, CNTECHTRA Research Branch Report, 1975.
- [28] Liss, M., Saulnier, C., Fein, D. and Kinsbourne, M. No Title. *Autism: the international journal of research and practice*, 10(2):155–72, 2006.
- [29] Lovaas, O. I. and Schreibman, L. Stimulus Overselectivity of Autistic Children in a Two Stimulus Situation. *Behavior Research and Therapy*, 9:305–310, 1971.
- [30] McNamara, D. S., Graesser, A. C., McCarthy, P. M., and Cai, Z. *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press, Cambridge, M.A., 2014.
- [31] National Autistic Society. A Place in Society: The Importance of Planning for Life for People with Autistic Spectrum Disorders. Technical report, 2004.
- [32] D. Neurosciences and L. D. Directorate. Exercise Is Good For You! Easy read leaflet, 2008.
- [33] O’Connor, I.M. and Klein, P.D. Exploration of Strategies for Facilitating the Reading Comprehension of High-Functioning Students with Autism Spectrum Disorders. *Journal of autism and developmental disorders*, 34(2), 2004.
- [34] N. Pavlov. User Interface for People with Autism Spectrum Disorders. 2014(February):128–134, 2014.
- [35] C. Putnam and L. Chong. Software and technologies designed for people with autism: What do users want? In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*, Assets ’08, pages 3–10, New York, NY, USA, 2008. ACM.
- [36] K. A. Quill. Instructional Considerations for Young Children with Autism: The Rationale for Visually Cued Instruction. *Journal of Autism and Developmental Disorders*, 27(6), 1997.
- [37] Sampath, Harini, Sivaswamy, Joyanthi, Indurkha, Bipin. Assistive Systems for Children with Dyslexia and Autism. In *SIGACCESS Newsletter*, number 96, pages 32–36, 2010.
- [38] Sasson, N. J. and Elison, J. T. Eye Tracking Young Children with Autism. *Journal of Visual Experiments*, 61, 2012.
- [39] J. Theeuwes, A. Belopolsky, and C. N. Olivers. Interactions between working memory, attention and eye movements. *Acta Psychologica*, 132(2):106–114, Oct. 2009.
- [40] B. Tronbacke. Guidelines for Easy-to-Read Materials. Technical report, IFLA, The Hague, 1997.
- [41] Vajjala, S. and Meurers, D. On improving the accuracy of readability classification using insights from second language acquisition. In *In Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, pages 163–173, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.