



A Preliminary Study of the Eye-Gaze Patterns and Reading Comprehension Skill of Students on the Autism Spectrum

Bradley M. Drysdale¹ · Brett E. Furlonger¹ · Angelika Anderson² · Dennis W. Moore¹

Accepted: 7 February 2022 / Published online: 16 February 2022
© The Author(s) 2022

Abstract

Objectives To identify, using eye-tracking technology, if readers on the autism spectrum (AS) would exhibit longer mean fixation times during question-answering but similar levels of comprehension accuracy to typically developing (TD) individuals.

Methods The eye-gaze behavior of two school-aged autistic children and two TD children with similar reading abilities, age, gender, and grade level, was tracked while they read age-appropriate passages and answered related comprehension questions.

Results Both participants on the AS displayed longer mean fixations than TD participants during question-answering. Despite the longer fixations during question-answering, the accuracy of the participants on the AS was equal to or superior to those of the TD participants.

Conclusions Longer mean fixation durations during question-answering did not result in lower reading comprehension scores.

Keywords Autism · Eye-tracking · Reading · Comprehension

The results of earlier studies led to the assumption that individuals on the autism spectrum (AS) answered reading comprehension questions less accurately because they exhibited longer eye-gaze fixations during text reading than typically developing (TD) individuals. This finding led the same researchers to conclude that text processing was more demanding and difficult for those on the AS than it was for TD individuals (e.g. Minshew & Goldstein, 1998; Sansosti et al., 2013). This conclusion may have been influenced by the results from studies of readers with dyslexia in which longer mean fixation durations and reading difficulties were to be correlated (Rayner et al., 2006).

In contrast, later studies indicated that, despite longer mean fixation durations during reading, individuals on the AS were observed to have been as accurate in answering reading comprehension questions as TD individuals (Howard et al., 2016; Micai et al., 2017) when matched on word reading and language skills. For example, Howard et al. (2016) compared the lexical, syntactic, and semantic

processing during the reading of both individuals on the AS and TD individuals and found individuals on the AS to have longer reading times overall yet similar accuracy to TD participants on measures of comprehension. Unlike earlier researchers (e.g. Minshew & Goldstein, 1998; Sansosti et al., 2013), who assumed that longer mean fixations were associated with more difficult and demanding processing, Howard et al. (2016) proposed that longer mean fixations to text reflected a “cautious” style of reading rather than difficulty comprehending the text itself which led to the conclusion that this behavior was characteristic of readers on the AS. Micai et al. (2017) confirmed the Howard et al. (2016) finding that both autistic and TD individuals were as accurate at answering comprehension questions even though the participants on the AS displayed longer fixations and more regressions to the target word during reading. More recently, non-invasive eye-tracking has been used in these later studies to examine eye-gaze movements by providing accurate, moment-to-moment information (Frank et al., 2012; Klin et al., 2002; Wilkinson & Mitchell, 2014) compared to earlier studies when eye-tracking involved attaching equipment to the user’s head, restricting their head movement.

Significantly, while eye-tracking has been used to assess eye-gaze during reading, it has not been used to measure eye-gaze behavior during the time when the individual is seeking the answer within text (hereafter referred to as

✉ Bradley M. Drysdale
bmdrysdale@gmail.com

¹ Faculty of Education, Monash University, Melbourne, VIC, Australia

² School of Psychology, The University of Waikato, Hamilton, New Zealand

question-answering): a potentially critical element missing from this area of research and unique to the present study. Although reading comprehension has often been defined as accurate text recall, without reference back to text, recent evidence suggests that reading comprehension is also determined by how the information is retrieved (i.e. eye-gaze during question-answering; Drysdale, 2021).

Accordingly, the present study was designed to investigate the eye-gaze behavior of students on the autism spectrum during question-answering to further an understanding of the possible links between eye-gaze and text comprehension. To this end, two research questions were formed. One, would participants on the AS exhibit longer mean fixation times during question-answering than TD individuals? Two, would participants on the AS display similar accuracy when answering comprehension questions to TD individuals?

Method

Participants

Following ethics approval and informed parental consent, participants were selected based on their scores on the Progressive Achievement Tests in Reading (PAT-R), an Australian-based test of reading comprehension and word knowledge (see <https://www.australiaeducation.info/Tests/K12-Standardized-Tests/reading.html>). Two students with a clinical diagnosis of autism were recruited having been diagnosed by either a psychologist, paediatrician, or both. Students who were considered to be typically developing and of similar age, grade level, gender, word reading, and language skills were included as a form of comparison. Word reading and language skills were measured by the Total Reading Composite Score on the Woodcock Reading Mastery Tests—Third Edition (WRMT-III; Woodcock, 2011) pre-assessment. The descriptive statistics for participants are presented in Table 1.

Procedures

Assessment

Five minutes of each assessment session was spent familiarising participants with eye-tracking equipment, followed by the 90-min WRMT-III assessment which was conducted to eliminate any participants with a severe reading deficit. Four participants met the inclusion criteria.

Eye-Tracking

During calibration and reading tasks, stimuli were presented by the laptop installed with Tobii Studio and viewed by participants via the connected monitor. Binocular eye movements were recorded remotely, meaning no contact was made between the participant and eye-tracker. Each participant sat in a chair with adjustable seat height to ensure the top of the monitor was aligned with the top of the participant's head. The chair was positioned approximately 65 cm away from the front of the monitor. Before starting the experimental task, participants undertook a five-point calibration procedure. This involved participants following a dot with their gaze as it moved throughout the screen and stopped at five locations.

Reading tasks were completed in a single session of 50 min duration. These sessions were video-recorded, allowing recordings to be re-watched if researchers failed to code answers in vivo. Participants were instructed to read the entire text out loud from start to finish. The Reading Phase immediately followed calibration via a button press from the researcher and began the moment the text stimuli were presented on the monitor. This phase ended when the participant finished reading the last word of the passage, which prompted the researcher, via button press, to switch to a screen containing a black background and white fixation cross to prevent re-reading occurring during the Reading Phase.

Table 1 Participant demographics and WRMT-III pre-testing scores

Student	Gender	Grade	Age (years: months)	Diagnosis	PAT-R scores	WRMT-III				
						Total Reading Composite Score	Basic skills	Reading comprehension	Listening comprehension	Oral reading fluency
JG	F	4	10:0	AS	130.6	109	121	106	103	105
SM	F	4	10:0	TD	130.4	107	118	102	85	104
JC	M	3	9:9	AS	137.1	102	110	115	112	98
MI	M	3	9:4	TD	132.6	109	118	113	100	97

Descriptive labels for SS of WRMT-III scale scores: well below average (≤ 69); below average (70–84); average (85–115); above average (116–130); well above average (≥ 131)

Participants were instructed to answer a series of questions provided by the researcher, referring back to the text if needed. To begin the Question-Answering Phase, the researcher switched to the text screen. Participants were encouraged with a verbal prompt to guess the answer if they indicated they were unsure, or had not responded within 1 min. The researcher ended the task via a final button press once all questions were answered.

Measures

All testing sessions occurred in a small room containing two chairs and a table designed to minimise distractions.

Woodcock Reading Mastery Tests—Third Edition (WRMT-III)

Form A from the WRMT-III was used to test Basic Skills, Reading Comprehension, Listening Comprehension, and Oral Reading Fluency leading to a composite score of Total Reading ability.

Eye-Tracking Measures

Eye-movement data was collected using the Tobii X2-30, a remote-based eye-tracker that utilises pupil centre/corneal reflection technique (Tobii Technology AB, 2014). The X2-30 is suited to young children who may move their heads during testing as it maintains tracking activity within a 20" × 14" area (width × height). The X2-30 was interfaced with a Dell laptop installed with Tobii Studio v 3.4.6, a software program used to control the depiction of text stimuli to participants via a 1080 resolution 24" monitor (19 × 20). The X2-30 was attached to the bottom of the viewing monitor via a magnetic bracket which held it in place. The front of the eye-tracker was positioned at a 45° angle so as to point directly at participants' eyes. A Logitech HD Pro Webcam C920 was also connected to the top of the viewing monitor to record the participants' faces during eye-tracking tasks.

Texts

Four texts were included, two from the Grades 3–4 level of the Wechsler Individual Achievement Test—Second Edition (WIAT-II; Wechsler, 2005), and two created by the research team. Created texts underwent proofreading by researchers, who evaluated each for its suitability for analysis. All text was presented in black monospace Century Gothic font (size 12) on white background with 1.5 line spacing, to ensure adequate space between lines so that ambiguous fixations (slightly above or below the text line) would not be wrongly attributed to a different line of text. Images were added to resemble WIAT-II texts.

Text difficulty was graded with the Flesch-Kincaid readability formula allowing created texts to be standardised and represent the reading ability of a 9 to 10-year-old child with 3 to 4 years of Australian education. The algorithm used to test grade level readability was $(0.39 \times \text{average number of words per sentence}) + (11.8 \times \text{average number of syllables per word}) - 15.59$. Word count ranged between 78 and 156 ($M = 122$, $SD = 34.1$). Text difficulty statistics for texts are presented in Table 2.

Questions

The existing questions for the WIAT-II texts were used (*Crickets* = 6, *Good Neighbours* = 5 questions) while questions were written for created texts (*Poppy the Puppy* = 9, *Camping Weekend* = 7 questions) to assess students' comprehension. While each text varied in the number of comprehension questions used, the total possible marks per text ranged from 10 to 12 (*Crickets* = 12, *Good Neighbours* = 10, *Poppy the Puppy* = 12, *Camping Weekend* = 12 marks). The comprehension questions were based on the taxonomy developed by Day and Park (2005), which can be used to develop comprehension questions for texts to help young students better understand what they read, rather than a hierarchical model such as Bloom's Taxonomy (Bloom, 1956), where lower-order thinking skills must be acquired before higher-order critical thinking skills. Question types included were as follows: Literal (10), refers to an understanding of the straightforward meaning of the text, such as dates, facts, vocabulary, and locations; Reorganise (8), based on a literal understanding of the text, students must use information from various parts of the text and combine them for additional meaning, and; Inference (8), answers are based on

Table 2 Text difficulty measures

	Crickets	Good Neighbours	Camping Weekend	Poppy the Puppy
Word Count	78	100	155	156
Paragraphs	2	6	3	9
Sentences	7	13	15	17
Av. sentences/paragraph	7	2.6	5	1.8
Av. word/sentence	11	7.6	10.3	9.1
Av. characters/word	4.3	3.8	3.8	4.1
Proportion passive sentences	0%	0%	0%	5%
Flesch Reading Ease ¹	92.3	94.9	89.9	82
Flesch-Kincaid Grade Level ^{2,3}	3.1	1.9	3.2	4.1

Note. Total word count for Grades 3–4 texts was 489

Ease of readability is indicated by ¹higher/²lower scores – ³as a US grade level

material that is in the text but not explicitly stated, involves combining literal understanding of the text with own knowledge and intuitions (Day & Park, 2005). There were a total of 26 questions across all texts. Describing foundational comprehension skills as “lower-level” has led educators to devalue this knowledge (Munzenmaier & Rubin, 2013), yet the purpose of the current study was to explore how students obtain this knowledge through the study of their eye-gaze patterns. Question formats included the following: forced-choice i.e. alternatives, true/false; open-answer; and closed-answer, such as fill-in-the-blank (cloze exercise). Before a question and satisfactory answers were assigned to a story, a consensus of suitability was required from the research team. Answers provided by participants were assigned a mark of 0, 1, or 2, with higher marks reflecting a fuller understanding of the text. For certain forced-choice format questions, a maximum mark of 1 was assigned for a correct answer. Examples of questions and answer criteria are provided in Table 3.

Data Analyses

Tobii Studio was set to filter out all fixations shorter than 80 ms as this was calculated to be an insufficient time for text processing, and longer than 800 ms, as this was assumed to indicate “mind wandering,” that is, when the reader continued to maintain eye-gaze on a fixed location without processing visual information (Smallwood et al., 2011). Participant eye-gaze was recorded for both phases, with the Question-Answering Phase also divided into scenes, a feature of Tobii Studio, which allowed a recording to be broken up into subsets of the overall recording. The number of scenes equalled the number of comprehension questions, for example, if a participant answered five questions, five scenes

were created for that passage. A scene began at the end of a question-utterance by the researcher and finished at the end of the reply-utterance by the participant.

Dependent variables were mean fixation duration and comprehension accuracy. Fixations were defined as the maintenance of eye-gaze for at least 80 ms at a single location; hence, mean fixation duration was calculated as the average length of fixations across texts for a single participant. Comprehension accuracy (Marks awarded/Total possible marks) scores were added and converted into an overall percentage for each participant.

Results

Both participants on the AS demonstrated longer mean fixation durations than the TD participants during question-answering. Despite the longer fixations during question-answering, the comprehension accuracy of the participants on the AS was equal to or superior to those of the TD participants.

Indeed, JG spent longer, on average, on each fixation (250 ms) to text during question-answering than SM (TD) (170 ms). JG made less fixations overall than SM (JG = 2815 fixations; SM = 3170 fixations) and spent a longer time overall in the Question-Answering Phase (JG = 812 ms; SM = 697 ms) than SM. Significantly, JG was equally accurate in answering comprehension questions (JG = 32/46 marks = 70%; SM = 32/46 marks = 70%) despite the difference in eye-gaze behavior.

The participant JC, also spent, on average, longer on each fixation (140 ms) to text during question-answering than participant MI (TD) (110 ms). JC also made fewer fixations overall than MI (JC = 3906 fixations; MI = 6679 fixations);

Table 3 Question and answer examples

Question type	Text	Question	Answer criterion
Forced choice (alternatives)	<i>Camping Weekend</i>	“Do you think this person is happy or sad to be at the camp ground?”	1: Happy 0: Sad/irrelevant answer
Forced choice (true/false)	<i>Camping Weekend</i>	“The children can do hiking, fishing, and kayaking at the camp ground.”	1: True 0: False
Cloze	<i>Poppy the Puppy</i>	“Poppy is white with _____.”	1: Black spots 0: Other answer
Short open answer	<i>Camping Weekend</i>	“What is this person’s favourite part of camping?”	2: Roasting marshmallows AND singing songs 1: Either roasting marshmallows OR singing songs 0: Irrelevant answer
Long open answer	<i>Poppy the Puppy</i>	“How would this story have been different if Poppy was a lazy puppy?”	2: Refer to both: Poppy would not have chased after the ball; Emily would have had to go get the ball 1: Refer to: Poppy wouldn’t have chased the ball (no reference to Emily) 0: None of the above points

however, he spent a shorter time overall in the Question-Answering Phase than did MI (JC = 822 ms; MI = 1087 ms). Interestingly, JC was more accurate in answering comprehension questions than MI (JC = 38/46 marks = 83%; MI = 34/46 marks = 74%).

Discussion

In this preliminary study, it was observed that while two school-aged children on the autism spectrum exhibited longer mean fixation durations during question-answering, the accuracy of their question-answering remained at least equal to that of the TD participants. Past research has reported associations between longer mean fixations when reading and poor comprehension accuracy (e.g. Minshew & Goldstein, 1998; Sansosti et al., 2013). In contrast, and acknowledging the small sample size of the current study, poor comprehension scores were not observed for participants on the AS who demonstrated longer average fixation durations in the Question-Answering Phase. This finding was in-line with that of Howard et al. (2016) and Micai et al. (2017) who concluded that the reading comprehension accuracy of students on the AS was facilitated by longer mean fixation durations and that such observed behavior may be characteristic of a more cautious search style.

Author Contribution BD: sought and obtained ethics approval, designed and executed the study, undertook data collection and analysis, and wrote the paper. BF: collaborated with the design and writing of the study, and editing of the final manuscript. AA: collaborated with the design of the study and assisted in analysing the data. DM: collaborated with the design and writing of the study, and editing of the final manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data Availability All data are relating to this project are available through Figshare.

Tobii Data Reading & Questions.doc.x(<https://doi.org/10.6084/m9.figshare.18132035>).Participant Reading Data.xlsx(<https://doi.org/10.6084/m9.figshare.18132008>).

Declarations

Ethics Approval The current study was approved by the Monash University Human Research Ethics Committee (Project Number: 1111), as it met the requirements of the *National Statement on Ethical Conduct in Human Research*.

Informed Consent All parents and their child (participants) provided informed consent to participate.

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals: Handbook I, Cognitive Domain*. Longmans.
- Day, R. R. & Park, J. S. (2005). Developing reading comprehension questions. *Reading in a Foreign Language*, 17(1), 60–73. <http://nflrc.hawaii.edu/rfl>
- Drysdale, B. M. (2021). *A preliminary study of search strategies of primary school students during reading and question-answering: An illustrative example using eye-tracking*. Unpublished Manuscript.
- Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy*, 17(4), 355–375. <https://doi.org/10.1111/j.1532-7078.2011.00086.x>
- Howard, P. L., Liversedge, S. P., & Benson, V. (2016). Benchmark eye movements effects during natural reading in autism spectrum disorder. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 109–127. <https://doi.org/10.1037/xlm0000289>
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, 59, 809–816. <https://doi.org/10.1001/archpsyc.59.9.809>
- Micai, M., Joseph, H., Vulchanova, M., & Saldaña, D. (2017). Strategies of readers with autism when responding to inferential questions: An eye-movement study. *Autism Research*, 10, 888–900. <https://doi.org/10.1002/aur.1731>
- Minshew, N. J., & Goldstein, G. (1998). Autism as a disorder of complex information processing. *Mental Retardation and Developmental Disabilities Research Reviews*, 4(2), 129–136. [https://doi.org/10.1002/\(SICI\)1098-2779\(1998\)4:2%3c129::AID-MRDD10%3e3.0.CO;2-X](https://doi.org/10.1002/(SICI)1098-2779(1998)4:2%3c129::AID-MRDD10%3e3.0.CO;2-X)
- Munzenmaier, C., & Rubin, N. (2013). Bloom's taxonomy: What's old is new again. Perspectives. *The eLearning Guild*.
- Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflection of comprehension processes in reading. *Scientific Studies of Reading*, 10(3), 241–255. https://doi.org/10.1207/s1532799xssr1003_3
- Sansosti, F. J., Was, C., Rawson, K. A., & Remaklus, B. L. (2013). Eye movements during processing of text requiring bridging inferences in adolescents with higher functioning autism spectrum disorders: A preliminary investigation. *Research in Autism Spectrum Disorders*, 7, 1535–1542. <https://doi.org/10.1016/j.rasd.2013.09.001>
- Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., Carlson, J. M., & Schooler, J. W. (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *PLoS one*, 6(3), e18298.
- Tobii Technology AB (2014). *Tobii X2–30 Eye Tracker User's Manual*.

- Wechsler, D. (2005). *Wechsler Individual Achievement Test – Second Edition (WIAT-II)*. The Psychological Corp.
- Wilkinson, K. M., & Mitchell, T. (2014). Eye tracking research to answer questions about augmentative and alternative communication assessment and intervention. *Augmentative and Alternative Communication*, *30*(2), 106–119. <https://doi.org/10.3109/07434618.2014.904435>
- Woodcock, R. W. (2011). *Woodcock reading mastery tests: WRMT-III*. Pearson.
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.