Dyslexia and AI: The Use of Artificial Intelligence to Identify and Create Fonts to Improve Reading Ability of Individuals with Dyslexia

Beverly Boals Gilbert
Education and Behavioral Science
Arkansas State University
United States
bboals@astate.edu

Allen Hays
Education and Behavioral Science
Arkansas State University
United States
jhays@astate.edu

Xiuzhen Huang
Computer Science
Arkansas State University
United States
xhuang@astate.edu

Ariane Pait
Communication Disorders
Arkansas State University
United States
apait@astate.edu

Leslie Parker
Graphic Design
Arkansas State University
United States
lmoore@astate.edu

Jake Qualls
Computer Science
Arkansas State University
United States
jqualls@astate.edu

Jonathan Stubblefield
Arkansas Biosciences Institute / Computer Science
Arkansas State University
United States
jstubblefield@astate.edu

Karen Yanowitz
Psychology
Arkansas State University
United States
kyanowit@astate.edu

Eli Richmond

Computer Science
Arkansas State University
United States
eli.richmond@smail.astate.edu

Tamra Washington
Communication Disorders
Arkansas State University
United States
tamra.washingt@smail.astate.edu

Abstract: Dyslexia, also known as a specific reading disability, manifests as a difficulty with reading and spelling. Dyslexia is a neurological condition caused by differences in how the brain processes information. Although exact numbers are unknown, dyslexia is estimated to affect between 5-17% of the U.S. population and 5–10% of the world population. Dyslexia has a significant negative impact on patients' ability to perform in an educational environment. This has further negative impacts on children's self-efficacy and willingness to read. Research suggests that different fonts may have an impact on reading ability of students with dyslexia, however information on the specific fonts and font characteristics that are beneficial are contradictory. This project seeks to use artificial intelligence to design a font with the ability to help people with dyslexia read. To do this, we will expose students with dyslexia to a variety of fonts with different characteristics and evaluate their reading performance. Using an artificial neural network, we will develop a model capable of predicting a font's readability. Finally, we will use a generative adversarial network (GAN) to generate a novel font. We will modify this approach by constraining the GAN to generate fonts that score as highly readable by people with dyslexia as predicted by the neural network trained on our data. Success in this project will positively impact the lives of people with dyslexia and improve educational outcomes.

Background and Significance

Dyslexia, also known as a specific reading disability, is defined as reading skills 1.5 or more standard deviations below IQ (Hamilton, 2020). The learning disorder appears primarily as a difficulty with reading and spelling due to a defect in the brain's processing of graphic symbols. Dyslexia can also be thought of as a learning disability caused by a neurological condition that alters the way the brain processes written material. Dyslexia is a neurological condition caused by a different wiring of the brain. There is no cure for dyslexia (Hamilton, 2020). Research indicates that dyslexia has no relationship to intelligence. Individuals with dyslexia are neither more nor less intelligent than the general population (Hamilton, 2020). Dyslexia is the most common cause of reading, writing, and spelling difficulties (Hamilton, 2020). It affects individuals with different ethnic, socio-economic backgrounds equally as well as males and females equally (International Dyslexia Association, 2020). Although exact numbers are unknown, dyslexia is estimated to affect between 5% to 17% of the U.S. population (Hamilton, 2020). The Yale Center for Dyslexia identifies dyslexia as the most common of all neuro-cognitive disorders with 20% of the population suffering from the condition (Yale Dyslexia, n.d.). Snowling (2000) also addressed the population of students who fail to become independent and effective readers in spite of normal intelligence identifying the cause as developmental dyslexia affecting 5–17% of children in lower elementary grades or primary schools.

Although there is no cure for dyslexia (Hamilton, 2020), researchers have suggested that using specific fonts may be an effective an intervention strategy for individuals with dyslexia to facilitate success in reading. Multiple fonts have been designed and marketed specifically to assist readers with dyslexia. These include OpenDyslexic and Dyslexie (Wery & Diliberto, 2017)(Marinus et al.,2016). However, empirical studies of these fonts have not borne out any benefit to dyslexic readers from the use of these fonts. Wery & Diliberto (2017) compared the OpenDyslexic font to the commonly used Arial and Times New Roman fonts without finding any improvement to reading speed or accuracy (Wery & Diliberto, 2017). Similarly, Kuster et. al., (2018) studied the proprietary Dyslexie font, comparing it to both Arial and Times New Roman without finding improvements to reading speed or accuracy. Marinus and colleagues (2016) also studied the Dyslexie font. They found that the

Dyslexie font did improve reading speed. However, the effect disappeared when the within-word spacing and between word spacing of Arial was made to match that of Dyslexie (Marinus et. al., 2016).

Results from studies examining specific font characteristics has produced conflicting findings. French et. al., (2013) performed a study demonstrating increased information retention with a disfluent, or subjectively difficult to read, font compared to Arial. The size of this effect was larger for participants with dyslexia than those without. However, when Rello & Baeza-Yates (2013) studied how reading speed in people with dyslexia was affected by a variety of fonts, they found that sans-serif, monospaced, and roman fonts were superior. These do not match the characteristics of the monotype corsova font used as a disfluent font in the French et. al., (2013) study.

The impact of improving reading fluency cannot be overstated in the 21st century. As children become more fluent readers, they may come to enjoy reading, which then leads them to be more likely to choose to *continue* to read, which then may lead them to be more fluent readers, and so on (Abbot et. al., 2017). Reading comprehension and desire to read may have long lasting impacts. For instance, many secondary schools use texts as a primary source of course material (O'Brien et. al., 2009). Performance in school then is correlated with future post-secondary activities and potential career choice and success. In 2018, only 58% of 18–24 year-olds in Arkansas received a bachelor's degree, one of the lower percentages in the country (National Science Foundation, 2020).

Approach

Aim 1: Develop a Novel Font for Dyslexic Readers

To develop a novel font for dyslexic readers, we will use a machine learning approach. We will collect performance data from dyslexic readers using a variety of different fonts. Following this, we will analyze the fonts and performance data to build a convolutional neural network regressor capable of predicting reading performance of dyslexic individuals using a visual representation of the font as an input. After training this model, it will assist in designing a font that produces the maximum predicted performance in dyslexic readers.

Part 1: Collecting Performance Data from Dyslexic Readers on Different Fonts

To collect a large amount of performance data, we have developed a portable computer application capable of automatically evaluating reading performance of a dyslexic reader on a selected font. The application selects one font from approximately 130 at random for each trial with a dyslexic reader. The application measures reading comprehension using objective measures. As each reader will be asked to perform the evaluation on multiple fonts to provide robust average measures for each font, the evaluation will be kept short and engaging to encourage full attention and participation from the reader. Each font will be given a final overall readability score based on the average performance of dyslexic readers when using this font. The DIBELS assessment, which is now available in the public domain and is an objective and accepted measure of reading performance, will be utilized for electronic presentation of fonts (DIBELS: University of Oregon, n.d.). Reading passages and sentences from the DIBELS assessment are used to ensure that reading levels are appropriate for the students. Establishing appropriate reading and vocabulary content is insured by modifying DIBELS.

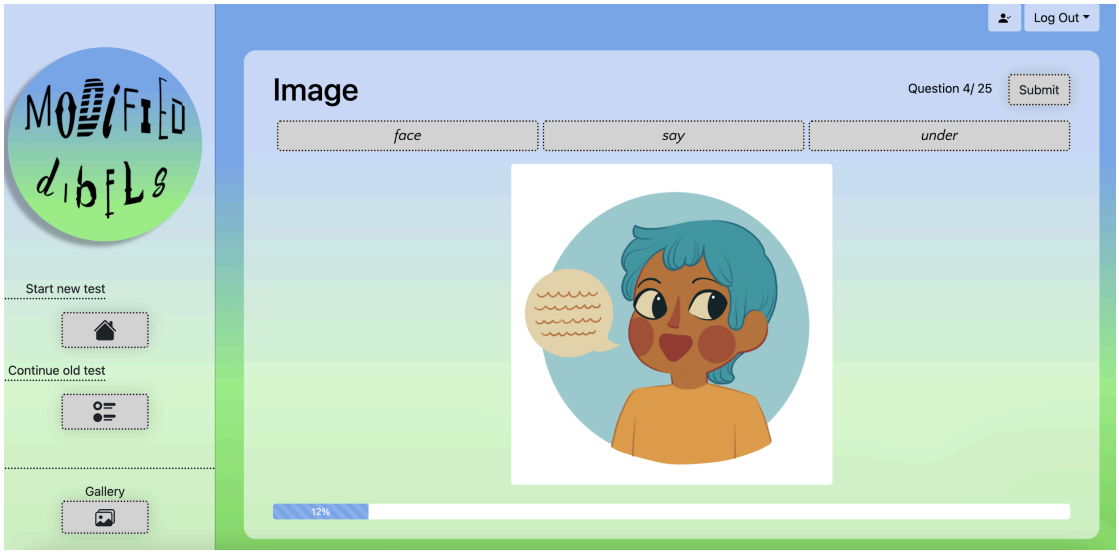

Figure 1: Example prompt from image task for grades K-1. Student must select the word corresponding to the image. Words will be presented in variable fonts.

Figure 2: Example prompt from maze task for grades above 1. Student must select the correct word to complete the sentence. Text will be presented in variable fonts.

Part 2: Building a Regressor Model for Readability

After collecting data and determining a readability score for each font, we will use the obtained data to train a convolutional neural network (CNN) to predict readability score given a visual representation of the font. CNNs are state of the art and standard practice for evaluating complex images including natural images, medical images, and handwriting. While training this model, we will use image augmentation to expand the apparent size of the dataset to the neural network and train the neural network to be robust against noise in the dataset. Before training begins, we will divide the dataset into training, testing, and validation partitions. We will estimate the error of the finished model using the reserved testing set.

Part 3: Using the Trained Model to Predict a Font with Maximum Readability

After a model has been trained that is capable of predicting the readability of an input font, we will use the model to design a font with maximum predicted readability. There are multiple methods by which this can be done.

A sample font can be modified incrementally until the readability of that font becomes maximized. Alternatively, the maximum predicted readability can be backpropagated through the neural network until an image with maximum predicted readability is output. Features of this image will be studied by an expert graphic designer and developed into a font. Similarly, the minimum predicted readability can be backpropagated through the neural network to produce an image with minimum predicted readability. An expert graphic designer can then develop a font that avoids all features of this image.

Current Project Status

Overview

A pilot study has been completed. More expansive data collection will begin now (Spring 2023) that approval has been received from the A-State Institutional Review Board to launch the full project

System Validation Pilot Study

A pilot study was completed to beta test the application. Target readers who consented to participate in the pilot were recruited from the Arkansas State University Speech and Hearing Center (ASTATE SHC) literacy intervention program during the Summer 2022 semester. Parents and students were invited to an information session prior to the start of the summer semester. They were informed of the pilot study, and any questions or concerns were addressed. Parents and guardians were not required to consent to have their child participate in the pilot in order to receive literacy intervention. Throughout the summer, students were provided with literacy intervention for four hours each week for seven weeks. On the day that they participated in the pilot study, students were escorted to a computer lab, assent was obtained, and the computer application was completed. Following the initial data collection day, students were asked about their experience interacting with the application. As a result, minor modifications in how the application looked were completed. No substantive changes were needed. It was discovered that the "submit" and "log out" icons were closer than researchers were comfortable with in order to avoid accidental premature ending of the application.

Results

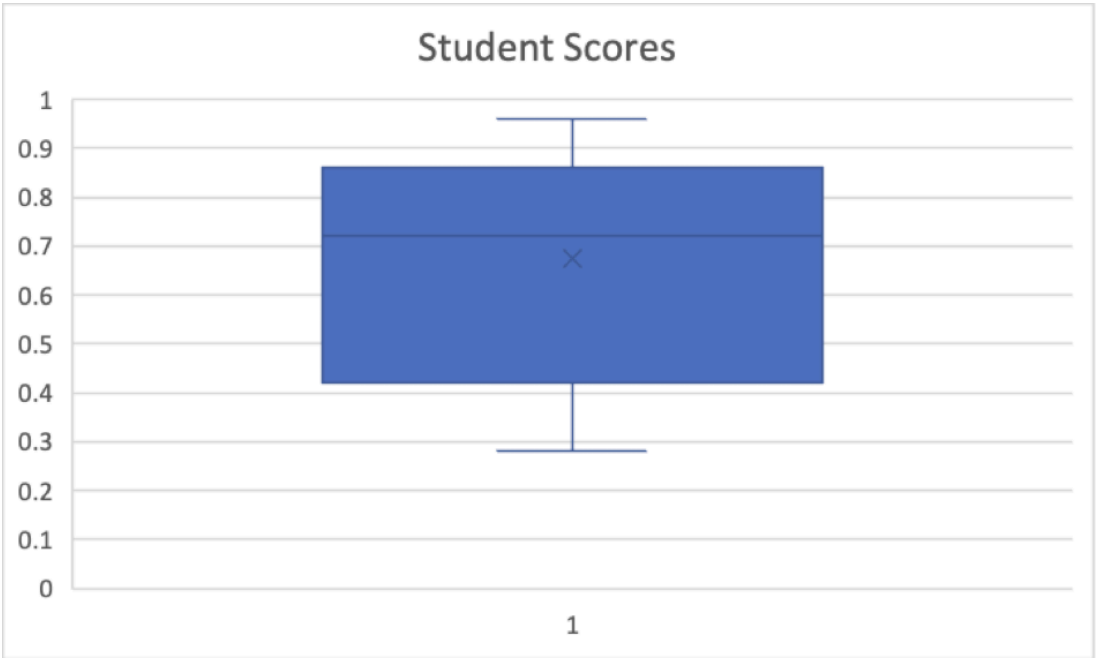

Figure 3: Box and whisker plot showing the distribution of student scores.

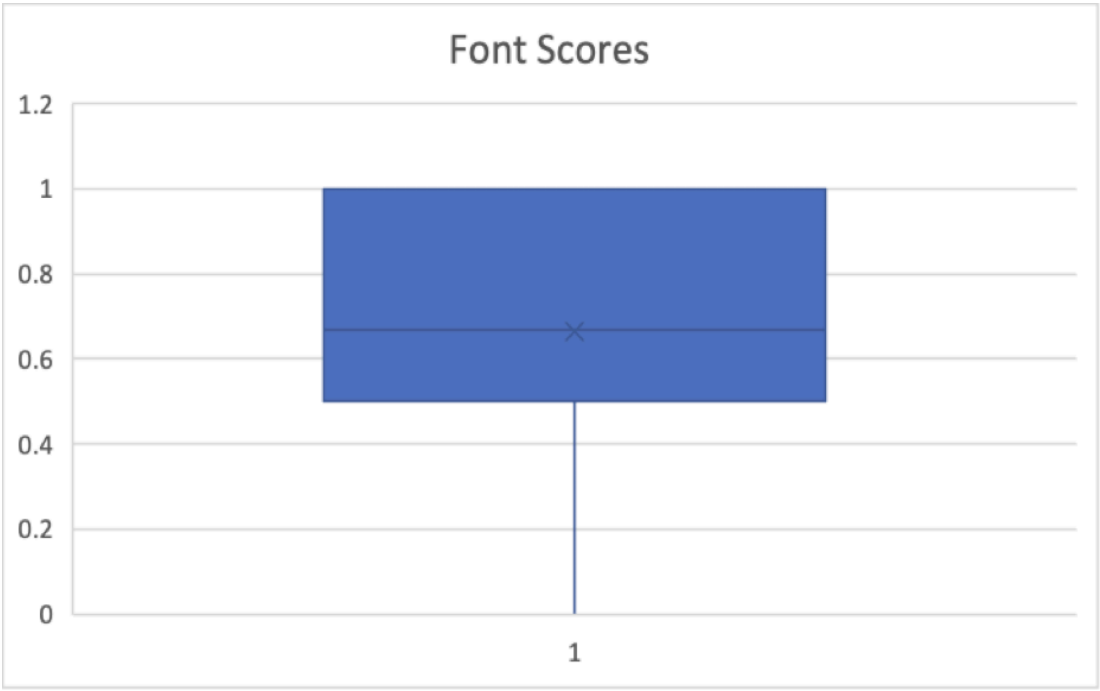

Figure 4: Box and whisker plot showing the distribution of font scores.

During the collection of preliminary data, 17 student trials were performed. Student scores during these trials averaged 65% accuracy. During these trials, the students were exposed to 132 different fonts taken from fonts for which attribute data is available. Each font was seen an average of 3 times with an average overall score of 66% accuracy.

Discussion

These results demonstrate that the data collection methods used are not inclined to introduce a bias by making the task too easy or too difficult for the students. Student scores approximate a normal distribution indicating students can perform the task while still leaving room for meaningful improvement. Font scores are distributed more widely. While the number of iterations for each font is small, this would indicate that certain fonts are more likely to result in higher student scores than others.

With these results, we can expect data from full deployment to be amenable to training a deep learning model capable of predicting font readability by children with dyslexia and predicting which font features are predictive of higher readability.

Next Steps and Future Directions

Full Deployment

Our focus for this project is to deploy in schools in the Mississippi River Delta of eastern Arkansas. The eastern region of the delta includes eight counties around eighteen school districts. The region has a high minority population and poverty rates compared with the rest of the state (U.S. Census Bureau, n.d.)

To deploy this project, we contacted two school districts within the region. These schools were selected due to ease of implementation, positive relationships between the study site and the school districts, willingness to participate, and low literacy scores. Our idea for implementation is to administer the test during the students' reading intervention pullout. The three schools that chose to participate in this research had low English Language Arts (ELA) scores in the 2021-22 school year. It should be noted that if a student scored "In Need of Support" or "Close to Ready," the schools require the students to attend reading intervention pullout. In school number one, 64.53% of

3rd to 6th-grade students scored “In Need of Support” or “Close to Ready.” In school number two, 90.38% of 3rd to 6th-grade students scored “In Need of Support” or “Close to Ready.” In school number three, 60.47% of students in 3rd to 6th-grade students scored “In Need of Support” or “Close to Ready” (Arkansas Department of Education, 2022).

Since the schools have high reading intervention student numbers, we can maximize data collection with students with reading difficulty without disrupting the education process. Principals have agreed to help us distribute permission slips to all parents of the reading intervention students. Permission forms will be distributed through school email as well as being handed out to parents during parent-teacher conferences. The data collection will occur during the students’ intervention. Students will receive a website link through the participatory schools’ Google Classroom, where the students can take the assessment. Data collection will take around fifteen minutes and will be collected once from each school.

Once data is collected from the selected elementary schools, we will contact the local educational cooperatives to meet with their literacy specialists. We will collaborate with the specialists to pick the schools we will approach for our second round of testing. We will approach the selected schools to seek their participation and proceed with the same method of recruiting and implementation as we did with the original schools.

Aim 2: Evaluate the Efficacy of the Novel Font for Improving Reading Ability in Dyslexic Readers

After our collection of novel fonts have been developed, their efficacy in readers with dyslexia must be confirmed. To do this, we will obtain cohort of dyslexic readers and randomly divide them into an experimental group for each novel font and a control group with a common font. These students will then be tested for their reading ability in each of the assigned fonts. Following the collection of this data, we will use an ANOVA, and if there are significant differences, conduct a series of pairwise post-hoc tests to compare average reading performance to determine effect size. Following this experiment, we will be able to identify our novel fonts as providing a significant improvement, significant detriment, or no significant change in reading performance over the control font.

Aim 3: Evaluate Student Engagement and Self-efficacy in Reading When Using Novel Fonts

Ultimately, the goal of many reading intervention programs is to improve reading comprehension and to encourage students to become more frequent readers. These factors are often interrelated, as students who read more frequently tend to have higher levels of comprehension (Cunningham & Stanovich, 1998). There are many factors that influence whether or not one child will read more frequently than another child, however, reading frequency and enjoyment is particularly low for children with reading difficulties, as they typically have to put more effort into reading but still encounter high levels of failure in reading tasks. This failure may lead them to be more reluctant to continue to engage in reading activities (Žolgar-Jerković, 2018). Difficulty in reading can have a cascading impact on school success, and then potentially career success. If the fonts developed in this project do succeed in increasing reading fluency, then we hypothesize that students will become more engaged readers as they become more successful at the process of reading. Surprisingly, a review of the literature by the PI/co-PIs reveal little research looking at the reading engagement of students with dyslexia.

Baker et. al., (2000) define engagement as “when they (students) read frequently for interest, enjoyment and learning. The heart of engagement is the desire to gain new knowledge of a topic, to follow the excitement of a narrative, to expand one’s experience through print…” (p. 2). Various components are important to engagement, including self-efficacy. Self-efficacy, or the belief that one’s actions will produce desired outcomes, is a crucial element in one’s decisions to engage in future actions (Bandura et. al., 2001). Several researchers have found the competence that a child believes s/he has in a task directly influences his/her academic and career choices (Jacobs & Eccles, 2000).

To assess Aim 3, students will be asked to indicate prior engagement and feelings of self-efficacy in terms of reading before engaging in the reading task and then anticipated feelings after reading texts. A randomly selected sub-set of students will also be asked qualitative questions to assess their feelings during the task and if they noticed any changes in font and/or reading fluency. Answers to the open-ended questions will be analyzed using content

analysis. The goal of these analyses will be to interpret participants' responses, by identifying potential themes that emerge from the data

Instruments

Reading Interest and Self-efficacy Inventory (R_ISEI, Abbot et. al., 2017). The R_ISEI uses participants' ratings on a Likert Scale in items such as "How likely are you to do reading that is not assigned?" and "Reading is hard for me" to assess self-efficacy.

Reading engagement survey (RES, Whitaker, 2009). The RES is designed to measure overall engagement. The survey is comprised of 18 items such as "It is important to me to be a good reader" and "I enjoy reading", using a Likert scale.

Aim 3 – Part 1

As described earlier, the first data collection from students (AIM 1) will ask students to read a variety of different passages in different fonts. Thus, the first part of examining the impact on engagement and self-efficacy will use a pre/post-test, paired t-test. Students will be asked to complete the scales before they start the reading task and then at the end of the task. Although results will not be tied to a specific font, this will reveal a general impact of using different fonts.

Aim 3 – Part 2

Aim 2 describes an experimental design that we will utilize here as well (comparing students' performance on control and novel fonts). Students will again receive the scales described above both pre and post reading task, and a repeated measures ANOVA will be used to analyze the results.

Future Work

Reading intervention programs aim to improve reading comprehension and encourage students to become more frequent readers. We will continue to strive towards this goal by extending this work by expanding data collection modalities to include mobile applications and incorporating eye tracking into the developed system for further evaluation of the reading process.

Acknowledgements

We would like to make special mention of the students in the 2021 Digital Illustration in the College of Fine Arts at Arkansas State University, who designed and created the child-friendly images used in the word-matching task as part of their curriculum.

References

Abbott, R., Mickail, T., Richards, T., Renninger, K.A., Hidi, S.E., Beers, S., & Berninger, V. (2017). Understanding interest and self-efficacy in the reading and writing of students with persisting specific learning disabilities during middle childhood and early adolescence. International Journal of Educational Methodology, 3, 41 - 64.

Arkansas Department of Education. (2022). *ADE My School Info*. Retrieved from https://myschoolinfo.arkansas.gov/

Baker, L., Dreher, M. J., and Guthrie, J. T. (Eds.). (2000). Engaging young readers: Promoting achievement and motivation. New York, NY: Guilford Press.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. Child Development, 72(1), 187–206. https://doi.org/10.1111/1467-8624.00273

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American educator*, *22*, 8-17.

DIBELS®. What is DIBELS? | DIBELS®. (n.d.). Retrieved November 11, 2022, from https://dibels.uoregon.edu/about-dibels

“Dyslexia At a Glance.” International Dyslexia Association, 7 July 2020, dyslexiaida.org/dyslexia-at-a-glance/#.

“Fact Sheets.” International Dyslexia Association, 4 June 2020, dyslexiaida.org/fact-sheets/.

M. M. J. French, Arabella Blood, N. D. Bright, Dez Futak, M. J. Grohmann, Alex Hasthorpe, John Heritage, Remy L. Poland, Simon Reece & Jennifer Tabor (2013) Changing Fonts in Education: How the Benefits Vary with Ability and Dyslexia, The Journal of Educational Research, 106:4, 301-304, DOI: 10.1080/00220671.2012.736430

Hamilton, S. S., MD. (2020, June 9). Reading difficulty in children: Normal reading development and etiology of reading difficulty (1142900557860264651 M. C. Patterson MD, FRACP, 1142900558 860264651 R. G. Voigt MD, FAAP, & 1142900559 860264651 M. M. Torchia MD, Eds.). Retrieved December 30, 2020, from UpToDate.

Hamilton, S. S., MD. (2020, September 8). Reading difficulty in children: Clinical features and evaluation (1142901253 860265065 M. C. Patterson MD, FRACP, 1142901254 860265065 R. G. Voigt MD, FAAP, & 1142901255860265065 M. M. Torchia, Eds.). Retrieved December 30, 2020, from UpToDate.

Jacobs, J.E., & Eccles, J.S (2000). Parents, task values, and real-life achievement. In C. Sansone, & J.M. Harackiewicz (Eds.), Intrinsic and extrinsic motivation: The search for optimal motivation and performance (pp. 405-439). San Diego, CA: Academic Press.

“Knowledge and Practice Standards for Teachers of Reading.” International Dyslexia Association, 18 June 2019, dyslexiaida.org/knowledge-and-practices/.

Kuster, S.M., van Weerdenburg, M., Gompel, M. et al. Dyslexie font does not benefit reading in children with or without dyslexia. Ann. of Dyslexia 68, 25–42 (2018). https://doi.org/10.1007/s11881-017-0154-6.

Marinus, E., Mostard, M., Segers, E., Schubert, T. M., Madelaine, A., and Wheldall, K. (2016) A Special Font for People with Dyslexia: Does it Work and, if so, why? Dyslexia, 22: 233– 244. doi: 10.1002/dys.1527.

National Science Board (2020). Science and Engineering Indicators 2020 State Data Tool. https://ncses.nsf.gov/indicators/states.

O'Brien, D., Stewart, R. & Beach, R. (2009). Proficient reading in school: Traditional paradigms and new textual landscapes. In L. Christenbury, R. Bomer & P. Smagorinsky (Eds.), Handbook of Adolescent Literacy Research, (pp. 80–97). New York: Guilford.

Para, J. (2021) Delta sees high rates of poverty. *Arkansas Democrat Gazette*. Retrieved from https://www.arkansasonline.com/1226deltapoverty/

Rello, L. and Baeza-Yates, R. 2013. Good fonts for dyslexia. In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13). Association for Computing Machinery, New York, NY, USA, Article 14, 1–8. DOI:https://doi.org/10.1145/2513383.2513447

Rosenzweig, E.Q., Wigfield, A., Gaspard, H., & Guthrie, J.T. (2018).How do perceptions of importance support from a reading intervention affect students' motivation, engagement, and comprehension? Journal of Research in Reading, 41, 625–641.

Snowling, M. J. (2000). Dyslexia (2nd ed.). Blackwell Publishing.

U.S. Census Bureau. (n.d.). *U.S. Census Bureau QuickFacts.* Retrieved from https://www.census.gov/quickfacts/fact/table/US/PST045222

What is dyslexia? Yale Dyslexia. (n.d.). Retrieved November 11, 2022, from https://dyslexia.yale.edu/dyslexia/what-is-dyslexia/

Wery, J.J., Diliberto, J.A. The effect of a specialized dyslexia font, OpenDyslexic, on reading rate and accuracy. Ann. of Dyslexia 67, 114–127 (2017).https://doi.org/10.1007/s11881-016-0127-1.

Whitaker, S.K. (2009). Development and validation of the reading engagement survey. Unpublished master thesis, A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree MASTER OF ARTS ATHENS, GEORGIA

Žolgar-Jerković, I., Jenko, N., & Lipec-Stopar, M. (2018). Affective factors and reading achievement in different groups of readers, International Journal of Special Education, 33, 201-211.