Web Tool based on Machine Learning for the Early Diagnosis of ASD through the Analysis of the Subject's Gaze

Sara Vecino[1], Martín Gonzalez-Rodriguez[1], Javier de Andres-Suarez[2] and Daniel Fernandez-Lanvin[1]
[1]Department of Computer Science, University of Oviedo, Spain
[2]Department of Accounting, University of Oviedo, Spain

Keywords: Eye Tracker, Automatic System, Diagnosis, Autism Spectrum Disorder, ASD, Machine Learning.

Abstract: Early autism spectrum disorder diagnosis is key to help children and their families sooner and thus, avoiding the high social and economic costs that would be produced other way. The aim of the project is to create a web application available in every health centre that could potentially be used as a previous step in early ASD diagnosis. It would be a fast way of diagnosing, spending almost no resources as it is web based. The system uses machine learning techniques to generate the diagnosis through the analysis of the data obtained from the eye tracker, and every time an evaluation is confirmed, it will be added to the training data set improving the evaluation process.

1 INTRODUCTION

Autism Spectrum Disorder (ASD) is a conduct disorder characterized by a person's intense concentration on their own inner world, as well as the progressive loss of contact with external reality. People with ASD can develop dependency on their family resulting in high social and economic cost, which may reach several million euros (Buescher et al., 2014).

The diagnosis of ASD is often based on tests and questionnaires where the data comes largely from observations of the child's behaviour done by relatives. Therefore, these data are subjective and may be inaccurate and/or lead to overdiagnosis (Davidovitch et al., 2021).

Recently, the efforts on ASD diagnosis moved to the use of objective observations mostly based on eye trackers, which capture the subject's gaze for a later analysis (Nag et al., 2020) (Shishido et al., 2019) (Cabibihan et al., 2017).

Average age of diagnosis of ASD coincides with the time of schooling, around 3 to 6 years of age. However, an earlier diagnosis would notably increase the effectiveness of the interventions (Øien et al., 2021).

This work pretends to provide an automatic tool to support an earlier diagnosis of ASD, by means of the analysis of the subject's gaze (using eye-tracking technologies) and machine learning (ML) classification systems.

The developed platform may provide support to multiple health institutions as the data gathered by their eye trackers as well as the diagnoses confirmed by their experts are shared in the cloud. This information feeds artificial intelligence-base agents who will provide an automatic diagnosis, based on the previous ones, for future subjects. As the knowledge based is increased with every new diagnosis, the accuracy of the computer assisted diagnosis should be increased too.

The resulting system will allow any paediatrician to get an automatic diagnosis of the subjects, detecting potential ASD children and anticipating their detection and treatment.

The remainder of the paper is structured as follows: Section 2 discusses previous works related to the topic of ASD automatic diagnosis. Section 3 describes the proposed approach, including the diagnosis procedure designed by the clinical experts collaborating in the project. Section 4 summarizes the results obtained in the preliminary study we conducted to evaluate the feasibility of this approach. Section 5 describes the architecture of the tool. Finally, Section 6 summarizes the conclusions, limitations, and future work.

Vecino, S., Gonzalez-Rodriguez, M., de Andres-Suarez, J. and Fernandez-Lanvin, D.
Web Tool based on Machine Learning for the Early Diagnosis of ASD through the Analysis of the Subject's Gaze.
DOI: 10.5220/0010715800003058
In *Proceedings of the 17th International Conference on Web Information Systems and Technologies (WEBIST 2021)*, pages 167-173
ISBN: 978-989-758-536-4; ISSN: 2184-3252

2 BACKGROUND

ASD is diagnosed around the age of 2 years old in the best case, when the ASD-related observed behaviours are more evident and so, the diagnosis is more trustworthy. The fact that most of the evidence of autism is related to interaction with other people delays the intervention of clinical experts until this age. Nevertheless, some evidence of ASD may appear at 12 months old or even earlier. For example, a typical sign is the lack of eye contact, that is, children with ASD find more difficult to maintain eye contact than children without ASD (Carette et al., 2019).

The analysis of eye gaze may lead to a solid and objective measurement and detection of evidence of ASD. The use of eye trackers has the advantage of being a non-invasive way of monitoring the children's gaze, revealing abnormal social behaviour without expending much time on the analysis.

However, there are many variables which can be measured. Among these, we can mention saccades, eye trajectories, bouncing tracks, time expended on the interactive elements (such as people, animals, etc.) and non-interactive ones (objects, background). These variables can be captured in a social interactive session (Cilia et al., 2018). The number of variables is so huge that it is difficult for humans to process such amount of information, in order to detect the abnormal social behaviour of ASD children (Saitovitch et al., 2013)..

2.1 Eye-tracking and Machine Learning Systems

There is a wide background of research about automatic ASD diagnosis. Most of prior efforts are based on the analysis of technical questionnaires like MCHAT that are fulfilled by clinical experts during subjects' evaluation. Hyde et al. analysed 45 papers that applied ML to ASD diagnosis (Hyde et al., 2019). Some other works explored the combination of eye-tracking and ML techniques for this goal. Jiang et al. (Jiang et al., 2019), for example, analysed the gaze movements of 23 subjects of an average age of 12,74 years old during facial recognition tasks. They built a Random Forests-based classifier obtaining an 86% of accuracy. Dris et al. (Dris et al., 2019) reached similar results combining eye-tracking analysis with Support Vector Machines (SVMs) classifiers. Wan et al. (Wan et al., 2019) tracked the gaze of a sample of 37 paired couples (ASD / non-ASD) while they view a video of a woman speaking for 10 seconds. They used SVMs and obtained promising results with children between 4 – 6 years old. Finally, Canavan et al. (Canavan et al., 2018) compared the performance of several ML algorithms (Random Forests, C4.5 and Partial Decision Trees [PART]) for the classification of ASD subjects, concluding that PART was the most effective for their specific approach. They used a sample of subjects up to 60 years old.

The results of previous works are not applicable to our specific context, since none of them worked with toddlers younger than 24 months old. In fact, most of the stimulus they used in their experiments are not applicable to children of these ages given that they still had not developed some of the interaction skills of an adult. However, the results of the combination of eye-tracking analysis with ML algorithms are promising in the context of ASD diagnosis.

3 OUR APPROACH

This project is developed in collaboration with ADANSI (A Spanish association of relatives and people with autism). ADANSI experts designed a set of videos focused on the stimulation of some reactions in the subjects that can be detected by the analysis of their gaze. In the first stage of the project, they implemented a manual procedure in which the subject's gaze was gathered using an eye tracker. Gaze information was processed manually, calculating some attributes that are not extracted directly from the eye tracker (for example, the number of exchanges between two different objects or persons in the video, among others). With this data, statistical analysis methods were applied to check the feasibility of the classification using computer-assisted methods. With the information of these attributes, the clinical expert can generate a diagnosis. This method was validated with currently accepted diagnosis procedures for ASD. Nevertheless, this is a lengthy process that takes up a lot of personal resources, not only for doing the tests with the children, that will remain the same, but also the time it takes to look at the relevant variables in each file and then decide if the child has ASD or not. The goal of this specific work is to automate all these tasks that are currently carried out by the expert. That is:

- Processing the raw data generated by the eye tracker.
- Identifying and filtering attributes.
- Using ML classification algorithms to automatically issue a diagnosis.

To automate the process, a web platform was developed, so that the person using the tool has only to upload the data from the eye tracker and receive the computer assisted diagnosis, without the intervention of a clinical expert. Thus, unspecialized paediatricians can identify potential ASD subjects and redirect the ones with strong evidence of ASD to an expert to get a confirmation of the diagnosis. This would be a great improvement as it could facilitate ASD identification in early stages, getting a preliminary accurate diagnosis.

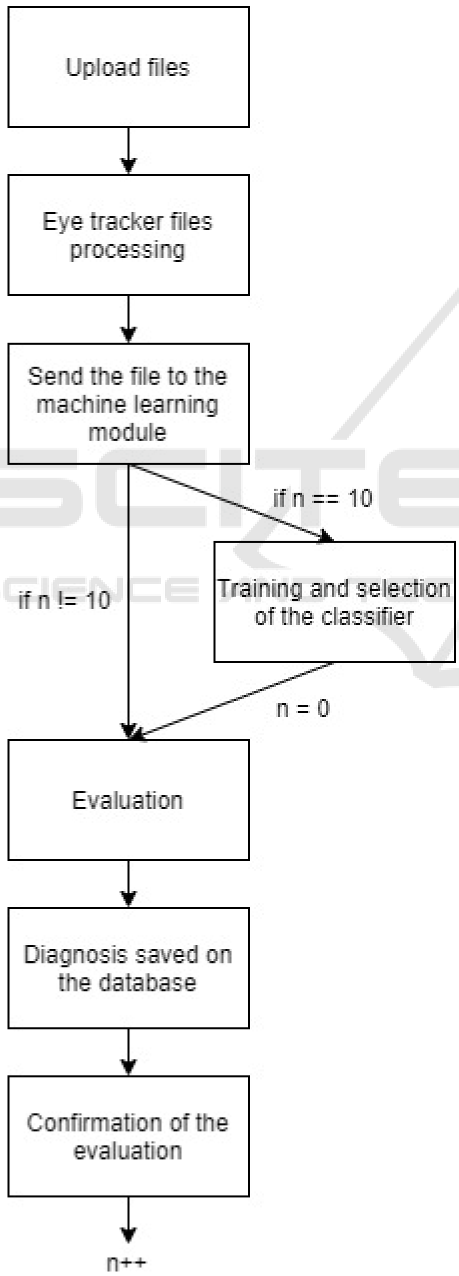

Figure 1: Diagnosis process. (Source: own elaboration).

The diagram shown in Figure 1 explains how the whole process works since the user starts uploading the files, after taking the tests with the eye tracker, until the diagnosis process has finished.

The process starts with the user uploading the files, the eye tracker files are processed as explained before and sent to the machine learning module. If it is the first time a user is automatically diagnosed or if the counter has reached the number 10, the system trains and selects a classifier to evaluate the user's data. If not, the evaluation is done directly with the currently selected classifier. Every time a new classifier is trained and selected, the counter is set to 0. Once the evaluation has finished, the diagnosis result is saved to the database and shown in the application. After that, an expert can confirm if the evaluation is correct or not, and the result of the expert diagnosis will be saved to the database as well. Once the subject diagnosis has been confirmed by the expert, the counter to retrain the system is increased.

In the following subsections, the different parts will be explained in more detail.

3.1 Visual Test Design

The process starts with an expert playing several videos to the babies while the eye tracker measures a wide spectrum of variables. These variables are later filtered by using correlation feature selection techniques to retain only the relevant ones. The initial set of videos and variables were selected by the ADANSI association as they are part of their research in early ASD detection.

The procedure established by the association consists of a total of eleven videos separated in three different categories:

- Social engagement.
- Social information gathering gaze exchanges.
- Gaze and deictic tracking.

3.1.1 Social Engagement

The aim of these videos is to analyse a child's response to a distracting element. An example of a video in this category is the one in Figure 2. It shows a woman singing a children's song while there is a cartoon dancing on the screen. The actress in the video tries to force a visual interaction between her eyes and mouth and the cartoon. The average social visual pattern for a situation like this implies to jumping the eye from the human to the character and back. A visual fixation over the character for a long period of time may be related to ASD.

Figure 2: Snapshot of a video in the category “Social engagement” (Source: ADANSI).

3.1.2 Social Information Gathering Gaze Exchanges

This group of videos is designed to check whether the children have difficulties for processing social information. For example, when an adult is referring to or pointing at an object (i.e., a toy). This is an important source of variables to detect ASD, as ASD-diagnosed children have been found to have problems when switching attention from humans interacting with an object to the object itself and back (Chawarska et al., 2012).

3.1.3 Gaze and Deictic Tracking

In this category, it is studied how children focus on several cartoons that appear on screen. The person in the video points or refers to the mentioned cartoons. These videos are designed in an increasing level of difficulty for the watcher. As the difficulty grows, the human watches to an empty space (making the eye to move towards it, and then to come back to the eyes of the human after a waiting time). An example of this type of videos is the one in Figure 3, which starts with a woman looking at her left when there is nothing on screen and then again when the cartoon appears.

Figure 3: Snapshot of a video in the category "Gaze and deictic tracking" (Source: ADANSI).

3.2 Data Filtering

The data files produced by the eye tracker are filtered to create two files, one containing the visual interchanges between elements (the human face and the cartoons) and another one with the number of fixations and time spent at each object.

3.2.1 Interchanges

This file contains the interchanges that the child makes between specific items and human body parts. An interchange can be defined using the following example. In the video corresponding to Figure 3, an interchange would be detected when the child looks at the woman’s eyes, then at the dog and then at the eyes again. This is represented by a series of ones and zeros in the file. If the child is looking at the eyes, the value is set at one, and in the other case it is set at zero. Interchanges cannot happen between inanimate objects; a body part must be involved.

3.2.2 Fixations and Time

In this file, the registered variables are:

- **Time to the first fixation:** The time it takes the child to fix the attention on the specific item is measured. The item can be the eyes, mouth, or another specific object in the video.
- **Total fixation duration:** The total time the fixation to a specific object lasts.
- **Number of fixations:** The number of fixations that the child makes to an item throughout the video.

All this data is processed from these files and organized into a final document that the machine learning module will analyse to generate the final diagnosis.

4 PRELIMINARY STUDY

Before the development of this system, a research effort was conducted to know if the diagnosis using machine learning techniques was viable or not. The study was done with a sample different from the one used for the development of the system. The variables of this additional dataset were extracted from a simpler set of videos. The sample was a balanced one with 37 pairs of subjects, meaning that for each child with ASD there was another one of the same age without it. In this study, the procedure followed consisted of several steps: 1) filter the variables that are most relevant to differentiate between subjects with ASD and without it, 2) train the ML models using the set of relevant variables, and 3) determine which is the best ML classifier.

The study was made using paired t-tests for the variable filtering, and different thresholds in the p-values: 0.05 and 0.1. We found that 0.05 consistently yielded better Area Under Curve (AUC) scores in the Receiver Operating Characteristic (ROC) curve for the subsequently estimated ML models. ROC-AUC is a commonly used statistic to assess the accuracy of ML classification models in the cases, such as ASD diagnosis, where misclassification costs are unknown. However, we decided to use the 0.1 threshold as less variables were removed from the original dataset and the scores do not significantly differ from the 0.05 ones.

Regarding the ML algorithms considered for the system, they were chosen after examining the results of the works reviewed in section 2. In previous research, it was evidenced that the most accurate models for ASD detection were decision trees, SVMs and neural networks.

Once the best suitable algorithms are selected, the sample was randomly split into training and test sets. This is made to avoid the risk of overfitting, which means that the system would perfectly predict the known data, but not new cases. In this case 80% of the sample goes to the training set, and 20% to the test set. This process was repeated 200 times, which means that for each ML model we had 200 estimation and test samples, and therefore 200 ROC-AUC scores.

The following results are those corresponding to the means of the ROC-AUC scores calculated for each algorithm using the procedure explained above:

P-value: 0.1

- Decision trees: 0.9094
- SVM: 0.9855
- Neural networks: 0.9783

P-value: 0.05

- Decision trees: 0.9018
- SVM: 1.0
- Neural networks: 0.9911

It is evidenced that the diagnosis can be made using the selected classifiers, as the means for the ROC-AUC scores are above 90% in all cases.

5 ARCHITECTURE OF THE SOLUTION

The automation of the explained process is implemented as a web application. It is a full stack application integrating three programming languages, JavaScript in the frontend, Java in the backend, and Python for the implementation of ML models.

The developed application can be divided into three different subsystems: the management module, the data processing module, and the ML module. The management module can be divided into the frontend and the backend of the application. The data processing module belongs to the backend, so it is a subsystem inside the backend part. The Spring Boot application is the centre of the application, to which the React application, the Python web service, and the database are connected. The system architecture is represented in Figure 4.

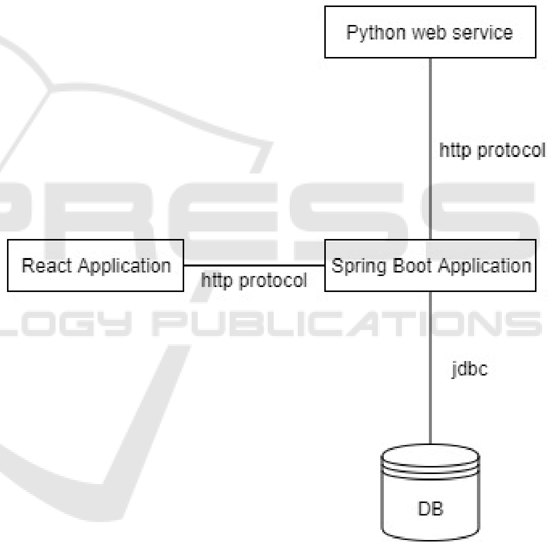

Figure 4: System architecture diagram. (Source: own elaboration).

This final system is built using a more complete and reviewed set of videos. As a result, the previous sample used to evaluate the feasibility of the study is outdated, and the current version is based on an unbalanced sample of 61 ASD / 2 non-ASD subjects. This unbalance will lead the system to provide bad initial classification accuracy, until new cases are added to the sample provided by ADANSI.

5.1 Eye Tracker Data Processing

The processing of the data obtained from the eye tracker is done in the backend of the application. This module receives the two-eye tracker produced files and transforms them into an information Data Transfer Object (DTO) that will later be processed in the ML module.

5.2 Diagnosis Engine

The aim of this module is to select the best classifier for each situation and apply it in the subsequent evaluations. The training of the ML models and the selection of the best classifier will be done every *n* new diagnoses, being *n* a configurable parameter initially set to 10. This means that there will be a first training and selection and after that, each time a subject is evaluated and confirmed by an expert, the new data will be added to the dataset. After doing this *n* times, the system will retrain the models and select the best classifier, which can be the same that is currently being used or a different one. This is important given that it is well known that classifiers' accuracy may significantly vary depending on sample size, and the variation in the accuracy of different algorithms when the size of the estimation sample increases may differ significantly.

For the decision trees, the split criterion used was the Gini index, and the split strategy was to choose the best split. The considered SVMs model used Radial Basis Functions (RBF). The neural network chosen was a multi-layer perceptron (MLP) classifier. The solver was the Limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS), an optimizer in the family of quasi-Newton methods. The neural network was configured with 3 hidden layers. All the models were implemented using the Scikit-Learn library for Python.

The ML module was developed as a Python web service, that connects to the main application using the http protocol. The backend of the application connects to the database using Java Database Connectivity (JDBC). To implement this solution, we chose PostgreSQL as this open-source database performs well with big datasets. It must be borne in mind that, although the sample considered in this study is initially small, the system is intended to grow as new cases (ASD and non-ASD children) are added.

5.2.1 Training Procedure

First, the dataset is filtered, to work only with the variables revealed as significant for the ASD / non-ASD classification. For this selection, as indicated in section 4, a paired t-test is used. Following the procedure set in the preliminary study, the threshold for the p-value is set at 0.1. If this threshold is exceeded the variable is not considered to be relevant for the diagnosis and therefore it is removed from the dataset. After filtering the data, the sample is recurrently divided into training and test sets in order to apply the cross-validation procedure. Then, the ML classifiers are created and fitted using the training sets. Then, ROC-AUCs are estimated using the test samples.

5.2.2 Selection

Using the aforementioned ROC-AUC scores for each classifier and each test sample, a t-test is performed to select which classifier is better for that specific case. In this case, an unpaired t test is used, as the process of generating the sets of estimation and test subsamples is separately conducted for each ML algorithm.

6 CONCLUSIONS

This tool is an automatic system to diagnose ASD in early ages, based on the data obtained from an eye tracker. The data is processed and evaluated using ML techniques. Out of the three considered ML algorithms: neural networks, SVMs machines and decision trees, the system chooses the best one while doing the training of the data, so that it is always adapted to the current sample. This system aims to help health professionals to diagnose ASD faster and easier.

As explained in section 5, the set of videos we used for the preliminary study to evaluate the feasibility of the classification system were extended and improved by ADANSI. In consequence, the current system has been developed with a new sample that, at the time of this publication, is incomplete and unbalanced. That means that the current classification results of the system may lack the required accuracy levels. However, the system's capacity to retrain itself every 10 new confirmed diagnoses will automatically solve this issue as soon as the sample is completed. Even more, the automatic selection of the best performing classifier prevents the system from degradation of the performance, given that some algorithms usually perform better than others with smaller training sets, while others are better when this set is big enough.

Finally, this system is still a work in progress and will be expanded in the future. Paired training will be implemented, so that the sample will always have a ASD subject paired with a non-ASD subject of the same age. It will also be considered that not all cases of ASD might be detected in the same way given that there are people with milder or more severe ASD-related behaviours and thus, diagnosed with different levels of ASD depending on the support they need.

ACKNOWLEDGEMENTS

This work was partially funded by the Department of Science, Innovation and Universities (Spain) under the National Program for Research, Development and Innovation (Project RTI2018-099235-B-I00) and by the Fundación Trapote (Ayuntamiento de Gijón)[1].

REFERENCES

Buescher, A. V. S., Cidav, Z., Knapp, M., & Mandell, D. S. (2014). Costs of Autism Spectrum Disorders in the United Kingdom and the United States. *JAMA Pediatrics*, *168*(8), 721. https://doi.org/10.1001/jamapediatrics.2014.210

Cabibihan, J. J., Javed, H., Aldosari, M., Frazier, T. W., & Elbashir, H. (2017). Sensing technologies for autism spectrum disorder screening and intervention. *Sensors (Switzerland)*, *17*(1). https://doi.org/10.3390/s17010046

Canavan, S., Chen, M., Chen, S., Valdez, R., Yaeger, M., Lin, H., & Yin, L. (2018). Combining gaze and demographic feature descriptors for autism classification. *Proceedings - International Conference on Image Processing, ICIP*, *2017-Septe*, 3750–3754. https://doi.org/10.1109/ICIP.2017.8296983

Carette, R., Elbattah, M., Dequen, G., Guerin, J.-L., & Cilia, F. (2019). *Visualization of Eye-Tracking Patterns in Autism Spectrum Disorder: Method and Dataset*. 248–253. https://doi.org/10.1109/icdim.2018.8846967

Chawarska, K., MacAri, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*(8), 903–913. https://doi.org/10.1111/j.1469-7610.2012.02538.x

Cilia, F., Garry, C., Brisson, J., & Vandromme, L. (2018). Joint attention and visual exploration of children with typical development and with ASD: Review of eye-tracking studies. In *Neuropsychiatrie de l'Enfance et de l'Adolescence* (Vol. 66, Issue 5). https://doi.org/10.1016/j.neurenf.2018.06.002

Davidovitch, M., Shmueli, D., Rotem, R. S., & Bloch, A. M. (2021). Diagnosis despite clinical ambiguity: physicians' perspectives on the rise in Autism Spectrum disorder incidence. *BMC Psychiatry*, *21*(1). https://doi.org/10.1186/s12888-021-03151-z

Dris, A. Bin, Alsalman, A., Al-Wabil, A., & Aldosari, M. (2019). *Intelligent Gaze-Based Screening System for Autism*. 1–5. https://doi.org/10.1109/cais.2019.8769452

Hyde, K. K., Novack, M. N., LaHaye, N., Parlett-Pelleriti, C., Anden, R., Dixon, D. R., & Linstead, E. (2019). Applications of Supervised Machine Learning in Autism Spectrum Disorder Research: a Review. In *Review Journal of Autism and Developmental Disorders* (Vol. 6, Issue 2, pp. 128–146). Springer New York LLC. https://doi.org/10.1007/s40489-019-00158-x

Jiang, M., Francis, S. M., Srishyla, D., Conelea, C., Zhao, Q., & Jacob, S. (2019). Classifying Individuals with ASD Through Facial Emotion Recognition and Eye-Tracking. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 6063–6068. https://doi.org/10.1109/EMBC.2019.8857005

Nag, A., Haber, N., Voss, C., Tamura, S., Daniels, J., Ma, J., Chiang, B., Ramachandran, S., Schwartz, J., Winograd, T., Feinstein, C., & Wall, D. P. (2020). Toward continuous social phenotyping: Analyzing gaze patterns in an emotion recognition task for children with autism through wearable smart glasses. *Journal of Medical Internet Research*, *22*(4). https://doi.org/10.2196/13810

Øien, R. A., Vivanti, G., & Robins, D. L. (2021). Editorial S.I: Early Identification in Autism Spectrum Disorders: The Present and Future, and Advances in Early Identification. *Journal of Autism and Developmental Disorders*, *51*(3). https://doi.org/10.1007/s10803-020-04860-2

Saitovitch, A., Bargiacchi, A., Chabane, N., Phillipe, A., Brunelle, F., Boddaert, N., Samson, Y., & Zilbovicius, M. (2013). Studying gaze abnormalities in autism: Which type of stimulus to use? *Open Journal of Psychiatry*, *03*(02). https://doi.org/10.4236/ojpsych.2013.32a006

Shishido, E., Ogawa, S., Miyata, S., Yamamoto, M., Inada, T., & Ozaki, N. (2019). Application of eye trackers for understanding mental disorders: Cases for schizophrenia and autism spectrum disorder. In *Neuropsychopharmacology Reports* (Vol. 39, Issue 2). https://doi.org/10.1002/npr2.12046

Wan, G., Kong, X., Sun, B., Yu, S., Tu, Y., Park, J., Lang, C., Koh, M., Wei, Z., Feng, Z., Lin, Y., & Kong, J. (2019). Applying Eye Tracking to Identify Autism Spectrum Disorder in Children. *Journal of Autism and Developmental Disorders*, *49*(1), 209–215. https://doi.org/10.1007/s10803-018-3690-y

[1] https://www.gijon.es/es/fundaciones/7