ALIZA: SMART MIRROR AS AUTISTIC EDUCATION ASSISTANT

2020-079

Project Proposal Report

Pavidha Lojini Rajendran

B.Sc. (Hons) Degree in Information Technology

Department of Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

February 2020

ALIZA: SMART MIRROR AS AUTISTIC EDUCATION ASSISTANT

2020-079

Project Proposal Report

B.Sc. (Hons) Degree in Information Technology

Department of Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

February 2020

Declaration of The Candidate & Supervisor

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
R.P.Lojini	IT17131216	

The supervisor/s should certify the proposal report with the following declaration.

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor: Date:

ABSTRACT

Autism is a neurodevelopmental syndrome that cause social communication and behavioral challenges that needs full-time care with therapies. According to WHO, one in 160 children has Autism Spectrum Disorders(ASD). This could be higher as the prevalence of ASD is far unknown in low and middle income countries. People with Autism has difficulty in understanding complex language, nonverbal communication as well as expressing their needs and emotions. Early intervention can speed up child's development and reduce symptoms of autism at earliest. Speech therapy plays key part in early intervention programs.

Autism kids have strong visual processing skills so using pictures and technology can be more effective than speaking. Technological advances can potentially lead to novel and more effective treatment strategies and enhance quality of life for people with ASD and their families [1]. The smart mirror proposed is known as "Aliza", have voice assistant in addition to its other core features, to interact with the user to express their need by choosing option and it provides simple activities to identify objects. It is important to keep talking to the kid about what's going on and encourage them with reinforcers. This can keep them more engaged with Aliza and interested to complete the activities. Aliza can create a significant impact in improving language and communication skills.

Keywords: ASD,Voice Assistant, Smart mirror

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

1.1 Background

Autism spectrum disorder (ASD) is a complex developmental condition that involves persistent challenges in social interaction, speech and nonverbal communication, and restricted/repetitive behaviors [2]. Autism is heterogeneous condition that not all the people with autism has same profile. Even the communication disability varies from person to person. About 40% of children with ASD do not talk at all. About 25%–30% of children with ASD have some words at 12 to 18 months of age and then lose them. Others might speak, but not until later in childhood [3].

No cure found for autism but early interventions can reduce few symptoms and improve their abilities such as speech therapy for language and communication challenges. Speech pathologists perform evaluation before designing the therapy program for the kid. One of the speech therapies is Alterative Augmentative Communication (AAC) which is based on pictures and technologies. Aided AAC, in particular, would seem well suited to individuals with autism because it has been argued that they process visual information more easily than auditory information [4]. Usage of technologies like iPad, iPod and computers has already started changing the lives of people with ASD. High-tech system like speech generating device prompt phrases in computer generated voice when they tap on a picture or word. These methods are more effective than just speaking.

In this research, the smart mirror includes voice assistant with simple speech activities. It is an enhanced version of speech generating devices. Aliza welcomes the user by prompting their name which is one of the function of the voice assistant and correlated with face recognition system of Aliza. It is important to keep talking to the kids about what they are doing and what is happening around the environment mostly during the learning activities. Whenever the kid feels bored and less engaged with the system, Aliza interrupt and prompt to them to swap to another activity. Along with that, creating conversational routine to prompt simple phrases can make an impact in their speaking skills. For example, saying "Get Set…" and waiting until they say "Go". Furthermore, Children with ASD needs to be appreciated for what they do. *"Praise is important to develop into a type of reinforcement because praise is a naturally occurring reinforcer"* [5]. One of the characteristics of autism is having a fixation on one single activity and the inability to change course without extreme agitation, it is helpful that

they know what to expect if they perform a task properly [6]. At the end of the completion of every tasks, voice assistant in Aliza will prompt positive reinforcers to encourage them.

In addition, Aliza provides picture based activities such as to find objects in a picture, repeat sounds and words. As mentioned earlier, speech generating devices just speak phrases of word or pictures that is tapped by the kid. But the proposed solution takes the input voice to evaluate whether the kid attempts to repeat the word or make sound. According to their response to the activity, progress level will be indicated. Through this system, it is expected to create a significant impact in language and communication skills of children with ASD.

1.2 Literature Review

Technology usage has prominently increased in treatment and research study of ASD. Existing researches were more like a mobile app, tablet based technology or speech generating device. Prospect, picture exchange communication system based instant messaging application [10] is about a mobile application where the user can send a simple message by combining cards to form a sentence from a series of cards with certain illustrations. The application allows children diagnosed with ASD to communicate using Instant Messaging (MIM) and at the same time providing a fun yet interactive learning experience.

Another research presents a game for kids with autism. This system is focused on giving a solution to verbal communication problems, which can give the users a useful way of communication and interaction with others [8]. Two applications have been developed in this system.one is to manage the application by parents, teachers to keep tracking of the kid's progress. Kid will be encouraged to be able to repeat the word spoken by the system. the voice signal will be analyzed to determine if it has been correctly pronouncing and the log of the results, user's data and progress fill be showed in management component.

Cara Wilson,Margot Brereton,Bernd Plodere and Laurianne Sitbon undertook their study over two autism primary school with minimally-verbal children aged 5 to 8.They have developed a very simple audio-visual dictionary app called MyWord,focused on building a personal collection of words.MyWord allows children to create a digital dictionary of words that mean something to them. The child has to initially take their own photos to insert into a dictionary format and then can add an audio recording of sound of word, enabling them to build personalized catalogue of words [9]. Authors have mentioned that children whom teachers had previously identified as finding it challenging to attend to class activities, such as learning sight words by wrote on the interactive whiteboard, were showing greater joint attention to the words in their MyWord app.

"Let's talk" is a voice output communication aid application for personal digital assistant for autistic children. By tapping a symbol on a screen of PDA with this application, a user can show his/her thoughts with pictures and sounds to others easily. It has 120 symbols based on daily life. There are 2 modes which can be switched depending on different situations of users. On "Supportive Mode", a supporter shows the application to a user. On Self-Use Mode, a user

can tell what he/she wants directly with categories and symbols. It is possible to make original icons with a camera or a voice recorder in PDA. A user also can customize an original page by arranging icons he made or existing symbols. The authors have operated a clinical observation to estimate the effect of this application to an autistic child [11].

In addition to the researches, there are number of mobile applications published and being used widely. Otsimo, Proloquo2 go are some of the existing mobile applications. Otsimo has developed more than 50 games to develop various skills. They have provided several games to improve communication skills. Kid can learn different words and sounds through this app. Proloquo2 is picture exchange communication system which is customizable. As per the researches mentioned above, there are technologies already prevailing to develop the communication skills but it lacks some of the features that can make the solution more suitable and easy to use.

1.3 Research Gap

Many researches are carried out to improve language and communication skills of children with ASD using technology. Technology based intervention is very effective in gaining attention and interest of the kid to focus on the tasks and skill development programs. Those researches have proven that it can produce positive outcomes. Augmentative Alternative Communication is a kind of assistive technology used by the speech pathologists. Mobile applications with AAC methods has already published. Most of these apps have enhanced the traditional way of Picture Exchange Communication System which is done using physical cards illustrating pictures. Along with PECS, computer generated voice read aloud the name of the card. It doesn't take the voice input of the kids to even check whether they have pronounced the word correctly whereas Aliza does.

Research A [10] uses series of card to send a simple message by combining them where the kid is not going to make an attempt to speak simple phrase or even repeat the word. Research B [8] is computer game to enhance the speech. This game has some of the functions that Aliza speech trainer provides. Compared to this game, Aliza doesn't have the feature to customize its activities. Research C [9] focused on building personal collection of words. Research D [11] is voice aid communication which is also fully customizable but lacking some other features like evaluating and logging the progress status. Below in table 1.1, Features of Aliza is compared with other existing researches and mobile applications.

	Progress Report	Effective reinforces	Voice input	Choice-based activities	PECS	Customization
Research A	✘	✘	✘	✘	✔	✘
Research B	✔	✘	✔	✘	✔	✔
Research C	✘	✘	✔	✘	✘	✘
Research D	✘	✘	✔	✘	✔	✔
Otsimo	✘	✘	✔	✔	✘	✘
Proloquo2go	✘	✘	✘	✘	✔	✔
Aliza	✔	✔	✔	✔	✔	✘

Table 1.1 – Comparison with existing researches

1.4 Research Problem

Worldwide one in 160 children has an ASD. This estimate represents an average figure, and reported prevalence varies substantially across studies. Some well-controlled studies have, however, reported figures that are substantially higher [12]. ASD is a range of conditions with developmental disabilities such as social and interaction challenges, repetitive behavior. Disability to communicate or understand what people say can create anxiety and stress to the children with ASD. Rapid development in technology has enhanced the learning process of them. Technological tool for intervention can ensure efficient learning and speed up the skill development. AI and modern technology is already applied in treatment and diagnose of Autism and it has created an impact in developing the skills.

Even though there many other technologies like humanoid robots AI apps are already deployed to the market, still some therapist or schools have not integrated them to their teaching process. Humanoid robots are more expensive to be used in all the schools and mobile apps lack key functions to attract the attention and motivate the children. The smart mirror "Aliza" with activities is a novel solution to develop various skills and it is proven that children with ASD are more engaged with mirror and likes seeing their reflection.

2. Objectives

2.1 Main Objectives

Main objective of this research is to improve the ability of language and communication of children with ASD.As previously mentioned, ASD varies person to person. Children with ASD who has limited speaking skills can benefit from the word repeating activities and recognizing objects and repeating sounds will at least make nonverbal kids try to speak.

2.2 Specific Objectives

1. To evaluate whether the kid repeats the word correctly or makes a try.

    Word recognition system matches the digital representation of the pre-defined word. Accuracy level may differ because of the time frame. It can be used to indicate the progress of the kid.

2. To prompt reinforcers after completing task successfully.

    After completing a task completely, positive words or visual effect or cartoon or music should be played as a reward to them.

3. To ask for swapping another activity when kid seems less engaged.

    Data of facial reaction recognition system grabbed from database, will be analyzed

3. Methodology

3.1 Data Collection

For training phase of word recognition system, word utterances should be collected from different ASD kids in audio file. It is planned to collect data from randomly selected nearly 100 kids in ASD pre-school in Colombo and Jaffna district. Voice recorder will be placed in child's pocket or close distance place to get recording by providing them word repeating activity and language skill development activities.
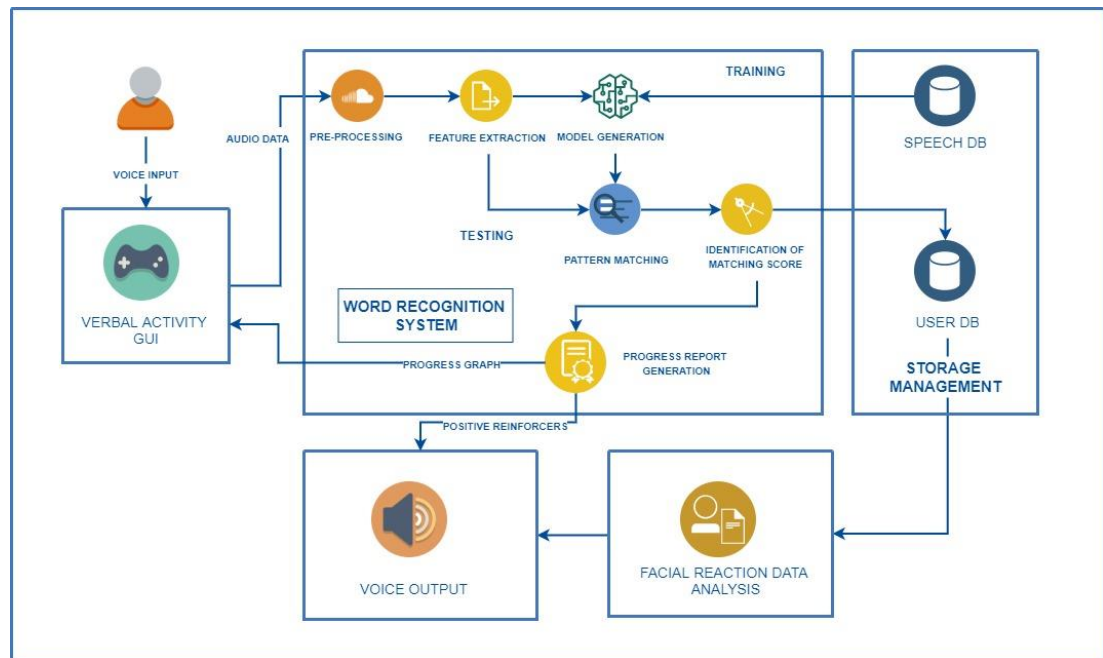
3.2 System Architecture

Figure 3.1 – High-level Architecture Diagram

One of the main components of "Aliza" system is voice assistant with language and communication skill development. Activities can be selected either by tapping on its icon or voice command. There are two types pf activities. One of the activities will be choice based where the kid has to tap on the correct picture for the question. The system will read aloud the repeatedly to make the child attempt to speak. Another activity is to tell the objects found in the given picture. It is important to keep the kids engaged with the activities to create an impact in their skill development. Considering the fact, Aliza includes an attentiveness tracker to monitor how they react during the activities. This is where attentiveness tracker and voice

assistant is correlated. Facial reaction processed data will be retrieved from the database and analyzed to prompt voice in between the activities as shown in figure 3.1. Output voice will raise a question, interrupting an activity to swap to another activity if the kid seems to be bored. In addition, Aliza encourages kids for completing tasks in the form of positive words or visual effects as a reward for them to perform well.
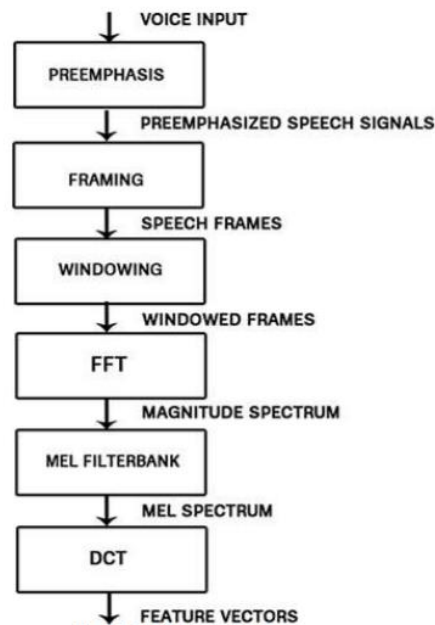
Furthermore, the core function of the sub system is word recognition system. As shown in figure 3.1, Audio data has to go through certain steps to produce the output. The steps to be followed are:

- **Audio pre-processing** –

    For training and testing, there would be collection of audio file taken from microphone or voice recorder with background noise. Preprocessing makes the processing of signal more convenient. So first noise will be removed from the recorded file.

- **Feature Extraction -**

    This is the most important step in this whole process since it will extract the relevant information from the speech frames. Audio signal needs to be converted to speech signal by a finite number of measures to be understand by the computer. There are different techniques available for feature extraction such as Spectrogram, MFCC and filter banks etc.

    MFCC (Mel-Frequency Cepstral Coefficients) can be used as feature extraction technique in this research. The numbers expressed in Mel Scale are more representative of human perception. Thus Mel Scale representation usually makes computers understand speech more

Figure 3.1 – MFCC extraction process [7]

like human, less like machines [7]. Feature extraction process in MFCC is shown in figure 3.2.

After converting the audio to signal, the way to recognize the word should be decided. Input audio file can be compared with prerecorded file by audio fingerprinting which is used in Shazam to find the music. Other alternative approaches such as using HMM (Hidden Markov Model) and ANN have to be researched to select the best fit method.

As the result of these steps, Confidence value is expected as the output to determine the progress of the kids.

4. Description of Personal and Facilities

Implementing a word recognition system to match with pre-defined words.

1. Preprocessing sound data
2. Match with pre-defined word representation
3. Analyze how accurate the word has been pronounced and generate progress level.

Analyze progress of other skill development activities and prompt positive reinforces.

- Grab data from other functions and provide positive reinforcers by phrases or visual effects.

Analyze data from facial reaction and prompt to swap to another activity.

- Grab data from facial reaction recognition system and when the system indicates the user seems to be bored, ask whether they wish to swap from the activity.
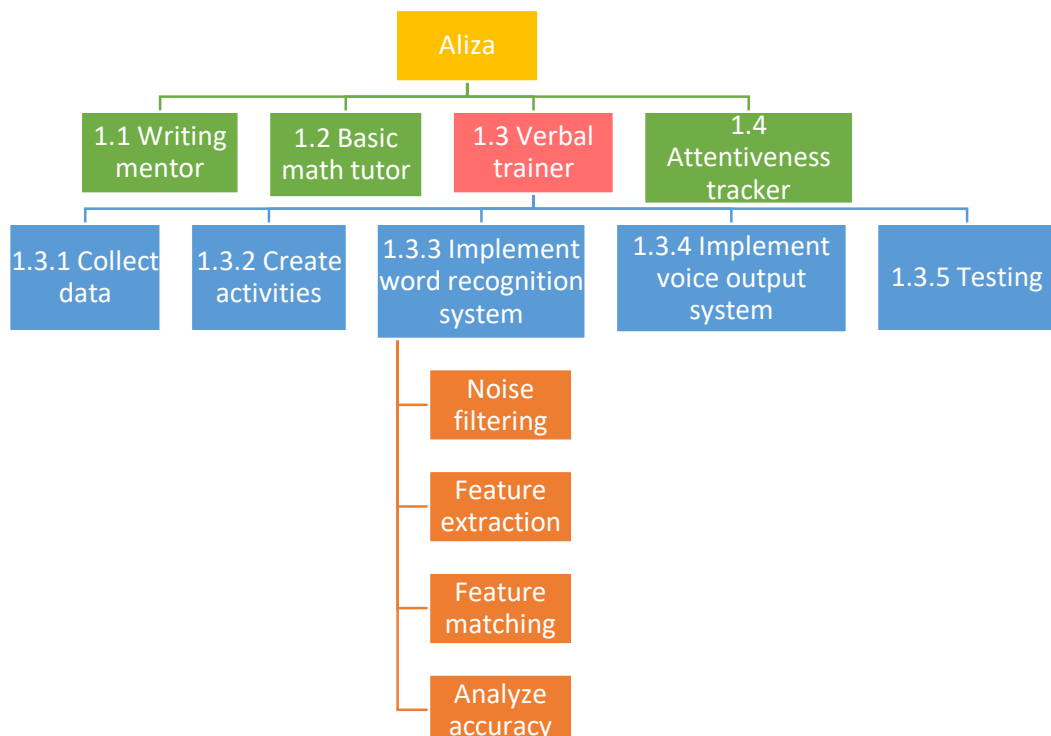
4.1 Work Breakdown Structure (WBS)

Figure 4.1 – Word Breakdown structure(WBS)

4.2 Self-Evaluation Plan

Evaluation Criteria	Excellent	Good	Average	Below Average
Gather required data	Collected nearly 100 audio recordings with best quality	Collected nearly 50 audio recordings with good quality	Collected nearly 20 audio recordings With average quality	Collected nearly 10 audio recordings With low quality
Build the smart mirror	Perfectly built the good looking mirror with all the required hardware	built the mirror with all the required hardware	built the with all the basic required hardware	built the good looking mirror with missing hardware requirements,
Implement the speech activities	Developed attractive game interface and well planned activities	Developed good looking game interface and some activities	Developed game interface and activities	Developed game with less user experience and incomplete activities
Build the word recognition system	Trained the model to acquire high accuracy and used best fit techniques	Trained the model to acquire moderate accuracy rate and good evaluation of other techniques.	Trained model with average accuracy rate	Not trained the model well and accuracy rate low
Testing the word recognition system	Correctly identified the word and high confidence value	identified the word and good confidence value	identified the word and average confidence value	Not identified the word
Retrieve data from facial recognition system and analyze it to invoke voice prompt	Prompted the voice exactly when the child seems less engaged	Prompted the voice exactly when the child seems less engaged	Prompted the voice exactly when the child seems less engaged	Missed to prompt voice

Table 4.1 – Self Evaluation Plan

5. Requirements

5.1 Functional Requirements

1) System should evaluate whether the user has repeated the word correctly.

2) System indicate progress of the user by analyzing how accurate the activity is performed.

3) If the user seems less engaged to the task, system ask to choose another activity to swap.

4) Prompting positive words to appreciate when an activity is performed well.

5) Begins with greeting the user by their name. (recognizing the user by facial recognition system).

5.2 Non-Functional Requirements

1) Usability

   Simple use of language and basic instructions provide for a user friendly system.

2) Affordability

   The system is at an affordable price that can be used for household as well as in schools. Moreover, the system facilitates many user accounts through one mirror.

3) Security

   The images of the individuals will only be processed and stored as a result. No individual data will be exposed.

6. Budget

Estimated budget can vary according to the decisions taken by the group throughout the project. Few of the components should be purchased online so its cost depends on the exchange rate of US Dollars. Given budget covers the hardware expenses of the product.

Component	Amount(LKR)
Wood Frame	5000
Two-way mirror	
Raspberry pi	8000
IR panel	20,000
LCD monitor	12,000
USB speaker	2000
Microphone	400
Power Supply unit	500
Camera	2500
SD card	300
TOTAL	50,700

Table 6.1 – Estimated Budget

7. Commercial Value

The commercial value of the proposed solution is increased through below mentioned information:

- The system accommodates many users through one mirror which can be used in schools and homes as well.
- The system provides parents with the facility to easier way to track their child's progress.
- Since the system is wall mounted smart mirror therefore easy to use and attracts many users.
- The system provides the facility to swap game activities by user as per their convenience. Especially this smart mirror has a user friendly voice assistant.

Reference

[1] Bölte, S., Golan, O., Goodwin, M. and Zwaigenbaum, L. (2010). What can innovative technologies do for Autism Spectrum Disorders?. *Autism*, 14(3), pp.155-159.

[2] Psychiatry.org. (2020). *What Is Autism Spectrum Disorder?*. [online] Available at: https://www.psychiatry.org/patients-families/autism/what-is-autism-spectrum-disorder [Accessed 20 Feb. 2020].

[3] Centers for Disease Control and Prevention. (2020). *Signs & Symptoms | Autism Spectrum Disorder (ASD) | NCBDDD | CDC*. [online] Available at: https://www.cdc.gov/ncbddd/autism/signs.html [Accessed 22 Feb. 2020].

[4] Iacono, T., Trembath, D. and Erickson, S. (2016). The role of augmentative and alternative communication for children with autism: current status and future trends. *Neuropsychiatric Disease and Treatment*, Volume 12, pp.2349-2361.

[5] Long, S. (2020). *Teaching Tip: Positive Reinforcement! - The Autism Helper*. [online] The Autism Helper. Available at: http://theautismhelper.com/teaching-tip-positive-reinforcement/ [Accessed 22 Feb. 2020].

[6] About Autism. (2020). *Positive Reinforcement - About Autism*. [online] Available at: http://aboutautism.net/positive-reinforcement/ [Accessed 22 Feb. 2020].

[7] Linkedin.com. (2020). *Speech Processing for Machine Learning. Spectrogram. MFCC*. [online] Available at: https://www.linkedin.com/pulse/speech-processing-machine-learning-spectrogram-mfcc-dipanjan-nag/ [Accessed 25 Feb. 2020].

[8] M. Frutos, I. Bustos, B. G. Zapirain and A. M. Zorrilla, "Computer game to learn and enhance speech problems for children with autism," 2011 16th International Conference on Computer Games (CGAMES), Louisville, KY, 2011, pp. 209-216.

[9] Wilson, Cara & Brereton, Margot & Ploderer, Bernd & Sitbon, Laurianne. (2018). MyWord: enhancing engagement, interaction and self-expression with minimally-verbal children on the autism spectrum through a personal audio-visual dictionary. 106-118.

[10] Raja, Pravind & Saringat, Mohd & Mustapha, Aida & Zainal, Abidah. (2017). Prospect: A Picture Exchange Communication System (PECS)-based Instant Messaging Application for Autism Spectrum Condition. IOP Conference Series: Materials Science and Engineering. 226.

[11] I. Torii, K. Ohtani, N. Shirahama, T. Niwa and N. Ishii, "Voice output communication aid application for personal digital assistant for autistic children," 2012 IEEE/ACIS 11th International Conference on Computer and Information Science, Shanghai, 2012, pp. 329-333.

[12] Who.int. (2020). *Autism spectrum disorders*. [online] Available at: https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders [Accessed 25 Feb. 2020].