Metrics Design of Usability and Behavior Analysis of a Human-Robot-Game Platform

HRG Metrics for LOLY-MIDI

Nayeth I. Solorzano Alcivar[1,2](✉), Luis C. Herrera Paltan[1,3], Leslie R. Lima Palacios[1,3], Jonathan S. Paillacho Corredores[1,4], and Dennys F. Paillacho Chiluiza[1,4]

[1] Escuela Superior Politécnica del Litoral, ESPOL Polytechnic University, ESPOL, Campus Gustavo Galindo, P.O. Box 09, 01-5863 Guayaquil, Ecuador
{nsolorza,lcherrer,ldlima,jspailla,dpaillac}@espol.edu.ec
[2] Facultad de Arte, Diseño y Comunicación Audiovisual – FADCOM, Guayaquil, Ecuador
[3] Facultad de Ciencias Sociales y Humanísticas – FSCH, Guayaquil, Ecuador
[4] Facultad de Ingeniería en Electricidad Y Computación - FIEC, CIDIS, Guayaquil, Ecuador

Abstract. As an innovative technological challenge, creating and designing metrics to evaluate communication between human-robot-game interaction will benefit children's education. In humans, facial expressions or emotions are pervasive forms of communication for interaction between people. When people are trying to establish communication deploying robots and game-based learning, which are growing in popularity, expectations are that these forms of relationship will become a means through which interaction is a common tool. Although it is intuitive for a regular human being to vary their expressions and emotions, their interpretation through metrics, or results of using the game as a form of learning, is a complex task that must be carried out. This paper explains the proposed design and usability metrics testing children's use of a human-robot-game platform, identified as LOLY-MIDI. This platform promotes inclusive education, primarily those children with Autism Spectrum Disorder (ASD).

Keywords: Human-robot-game · Metrics · Dashboard · Games-based learning · EDG · Autism · ASD · SAR

1 Introduction

The development of new technologies and the emergence of skillful generations in handling them is an increasing fact. The presence of smartphones and tablets in almost every home has awakened investigators and software developers' interest in implementing many applications as learning tools. Videogames or digital games built on educational properties, identified as educational digital games (EDG) or serious games, are an excellent way of teaching content and desired behavior, creating a good impression and growing acceptance on users for grabbing the attention of children with learning purposes [1, 2]. Social Assistive Robots (SAR) also contribute to the learning process and promote user interaction and active listening [3], expecting to motivate and capture

M. Botto-Tobar et al. (Eds.): ICAT 2020, CCIS 1388, pp. 164–178, 2021.
https://doi.org/10.1007/978-3-030-71503-8_13

their attention. In this sense, the LOLY-MIDI project is a Human-Robot-Game platform joining a social robot (named Loly) with an EDG series (named MIDI-AM) as mobile applications. All altogether seek to enhance children's learning experience initiating primary school, mainly aiming to capture and motivate children's attention with the Autism Spectrum Disorder (ASD) for inclusive education [4]. However, more than a lesson learned needs to consider confirming HRG platforms' influence on children learning and motivating their attention.

The SAR is a new subfield in human-robot interaction (HRI), focusing on developing intelligent robots to assist with social interaction. The SAR systems have great potential for providing personalized and affordable therapeutic interventions to children with ASD. However, there are limitations in the HRI that make interpretation and an instant response difficult when the robot is not in the presence of an operator. These limitations imply that the personalization of the SAR for each user is impossible in the actual context. Children with ASD have noticeable differences in language and child-to-child variations concerning cognitive ability [5], behavioral difficulties, and social understanding [6]. Due to this variation, individuals' learning styles across the spectrum are not uniform in nature [7].

ASD is a neurological and development affection commonly diagnosed during childhood, remains for a lifetime, and influences behavior, interaction, interpersonal communication, and the learning process [8]. For decades, studies have confirmed that children with ASD have a deficiency in their social-cognitive domain and are usually averse to participate in social interactions [9]. For example, this disorder affects one in 59 children in the United States [10] and 1266 persons in Ecuador [11]. Studies reveal that SAR has already been used to educate children with ASD to play with them to evaluate their contribution to these children's learning process. Besides, studies about Human-Computer-Interaction (HCI) point out that educational videogames help reduce the apathy and distraction they have as part of their characteristic behavior. [12, 13]. This investigation aims to determine the metrics that need to be used to measure the HRG applications (Loly Robot and MIDI-AM games) usability and their influence on children's learning experience and motivation, particularly in children with ASD.

1.1 The State of Art

Research carried out by social robotics has shown the importance of studying human-machine interaction to measure the degree of attention that a person maintains with digital games [14]. It is essential to study the individual's degree of attention to know if the robot's instructions reach the person. On the other hand, according to Gros [1], digital games can produce simplification of reality for promoting challenge, problem-solving, cooperation, and other motivations to encourage learning and cooperation. As a way of technology aids focusing their design on the education environment, these games are considered appropriate for improving learning [1, 2]. Thus, both resources educational digital games - EDG, and the socially assistive robot – SAR, have been recommended for educational purposes.

MIDI-AM is a research project based on the development of EDG in mobile applications for children. This project encourages the creation of Multimedia Interactive Didactic Infantile (MIDI) productions as free mobile applications. The MIDI-AM study is seeking to generate an objective evaluation of the impact and usability of games-based learning. In 2018 as part of this study, a dashboard module was designed for MIDI-AM games. This module retrieves data from the cloud generated when the games are played. The dashboard shows a preliminary statistical graph interpretation of the data about usability and playability of the MIDI-AM games [15, 16]. However, the MIDI dashboard results are not related to the effect of using these EDGs linked to the Loly SAR.

Considering there is little to no consistency in SAR's appearance used in therapy for ASD and their scarcity for therapeutic purposes in the market, the project Loly-MIDI was undertaken to accomplish these purposes. A robotic bust named Loly was developed to work linked to MIDI EDG applications. Loly is a non-anthropomorphic robot that represents an emblematic bird of Ecuador's coast region. Loly's primary function (head and bust) is to interact with regular children and ASD patients, helping them keep focused on a single activity [4]. This robot has a frontal camera and a 7-in. LCD screen to show different aye's expressions such as happiness, sadness, anger, and enthusiasm. For the verbal interaction, Loly uses a speaker through which it expresses pre-established lines of dialogues. Loly repeats the dialogues while the MIDI games are played, encouraging them to keep going or giving them instructions. Loly's wings have two degrees of freedom that it can move to emulate a flying bird through a small engine controlled by a microcontroller [4]. More studies with similar functionality, such as the robot named KIWI of Jain, Thiagarajan, Shi, Clabaugh, and Matarić [17], can be taken as a reference for the metric design to evaluate ASD children behavior using SAR.

It is argued that all social assistive robots look to generate one or more therapeutic interactions between humans and themselves, such as promoting a specific reaction, training skills, or simply serve as some sort of training for the user's social abilities linked to specific mobile applications [17]. Loly aims to play the role of a guide in the progress of different activities within the MIDI games, and at the same time, incentivize user's development of social skills [4]. Besides, it is relevant to highlight that, even with video games' incursion as game-based learning in the classrooms, still certain distrust from educators (Serrano-Laguna et al., 2017). Therefore, it is required to show convincing data that support the effectiveness of using SAR and EDG as linked platforms such as HRG proposed by the Loly-MIDI studies. Their contribution to learning and social skills development. In which new metrics can assist in evaluating the HRI and HCI according to each case's requirement.

For collecting data from the EDG series, a JavaScript Object Notation (JSON) record is required. The data is then stored in a cloud database to be further displayed in a dashboard [15].

A facial recognition application needs to be applied to analyze a SAR interaction with children, such as OpenFace [17]. There is some open free access to facial recognition software such as OpenFace. This application allows us to identify the user's facial landmarks and collect attention metrics to measure the degrees of attention and empathy during an interaction. OpenFace can provide training and test code, helping to replay the experiments of eyes and face recognition. It operates in real-time with all the modules involved in analyzing facial behavior, which is useful for this study [18].

2 Materials and Methods

2.1 Methodology

This research has a pragmatic perspective, using a qualitative-quantitative mix method strategy [19, 20]. A methodology for the design of an evaluation of LOLY-MIDI metrics separated into three stages is applied. The first stage begins by analyzing the interaction between the user with the mobile application and the robot. The second stage includes real-time data gathering, which happens simultaneously while the Loly robot is used with the game. The third stage occurs once the interaction has ended, and data has been saved, a final report is generated on a dashboard. The designed proposed metrics need to be used to generate the outcomes for evaluations and hypotheses that need to be confirmed or rejected to the last stage.

First Stage
Focus groups with regular children and with ASD are undertaken to observe and video recording users to interact directly with the robot and the mobile application. As a play, the users completed different playing exercises and heard different stories from the EDG. At the same time, children get feedback and instructions from the robot, who serve the purpose of a guide and the interaction.

Second Stage
OpenFace application oversees monitoring and gathering real-time facial data of the user. With this data, it is possible to analyze characteristics such as emotions, head pose, and eye gaze through pre-established conditions that interpret OpenFace's output metrics.

Third Stage
During the interaction, screen reports showing performance, attention, and emotions during the interaction are finally generated in a dashboard. With the expectation to implement this platform in primary schools of inclusive education, the operability and data presented on the dashboard must be accessible to a primary school teacher and psychologist. Therefore, simple easy-to-read graphs and statistical conclusions translated to natural language constructing simple statements are also requested. The use of simple language statements is relevant to engage with most potential users and minimize human errors in analyzing and interpreting the data.

2.2 LOLY-MIDI Beneficiaries

The direct beneficiaries for which LOLY-MIDI applications are intended are children between 3 to 7 years old, but mainly for children with ASD. Other beneficiaries are the parents, guardians, psychologists, and educators of primary schools. One of the significant challenges in this research has been to find volunteers that are eligible for the focus group experiment, despite working with different educative institutes that treat ASD in Guayaquil.

Children considered for this study should meet the following criteria:

a. The child with ASD should live in Guayaquil.

b. Should receive at least one session per week.
c. A teacher, parent, or guardian should accompany the child during the interaction with the LOLY-MIDI platform.
d. Parents and guards must sign a consent form for their child to participate in the experiment.

2.3 Research Tools Design

The robot prototype used is a parrot shaped robotic bust named Loly. The robotic bust has a camera that records a frontal view of the participants. Visual characteristics are obtained through the OpenFace software used to monitor emotions, head pose, and eye gaze. The data collected in real-time by OpenFace is intended to measure the user's degree of attention and empathy. To measure the user's attention level to the robot is expected to identify the segments on which the user's attention got lost and determine if there are significant differences in the degree of attention and empathy that Loly receives concerning the game played. Also, hypotheses are generated based on focus group experiments.

The newly proposed metrics are described as follows:

Face Detection – Emotions

There are a wide variety of available free tools that offer face detection in pictures or videos. However, few of those share their code freely. Most of them instead provide binary commands to use. Binary commands allow access to pre-established functionalities and often do not support different platforms, making their implementation almost impossible. These binary commands cause the application of facial detection tools to become a tricky task in experiments with different parameters and databases.

OpenFace applies Conditional Local Neural Fields (CLFN) [21] for facial landmark detection in real-time. CLNF is an instance of a Constrained Local Model (CLM) [22], which uses a better optimization function and more advanced expert patches. The CLNF model detects up to 68 facial landmark points (see Fig. 1). OpenFace uses this model and makes it better by training the distribution of facial landmark points in different groups (eyes, lips, and eyebrows).

Many binary variables are created from the OpenFace output that makes the identification of emotions possible. Emotions are determined based on a variation analysis of the distance between specific facial points identified and labeled by OpenFace. Distances referred to in the complementary material are tentative and will be determined with experimentation. The precision with which the distance between the facial landmark points is expected to work better is if the camera is at head level and correctly set up [24].

Head Pose

Unlike face detection, the head pose has not received the same attention from software developers. Many criteria allow for head pose estimation by using in-depth data [25]. However, they do not work appropriately on webcams. Some facial landmark detectors include the capacity of estimating head pose [26, 27], but again, most of them do not have support for webcams.

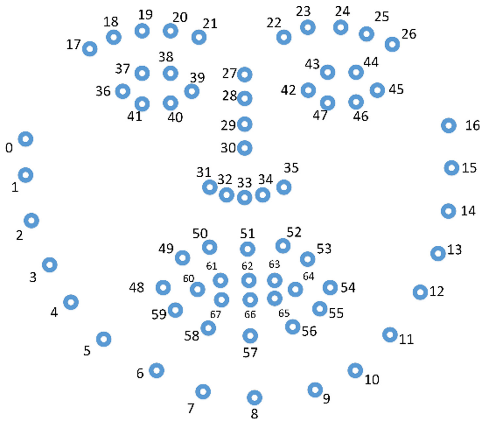

Fig. 1. OpenFace facial landmark points [23]

OpenFace can gather information about the head pose on top of facial landmark detection. These landmark detections are achieved because CLNF internally uses a 3D representation of facial landmark points, showing them on the image using an orthographic projection of the camera. The facial landmark points allow for head pose estimation once facial landmark points are correctly detected. OpenFace needs to have calibration parameters for the camera (focal distance and principal point). Without calibration parameters, OpenFace uses an estimation based on image size [18].

Binary variables are created from OpenFace's output to determine head pose. Two possible scenarios are considered, one in which the head is perpendicular and another one in which the head is tilted down (see Fig. 2). Angles proposed in the complementary material for binary variables are subject to changes in experimentation. They are expected to depend on the distance between the user's face and the camera and woks optimally if placed at head level [18].

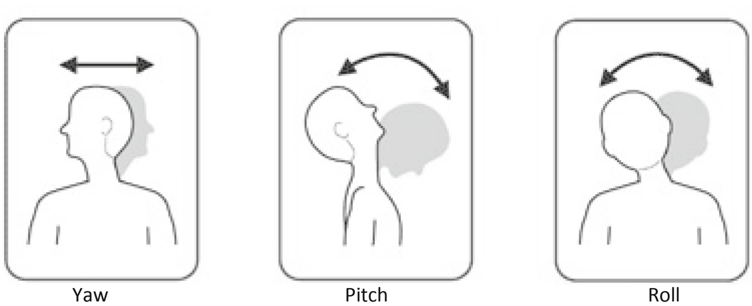

Fig. 2. OpenFace head pose [28]

Eye Gaze

There are numerous tools and commercial systems to estimate where someone is looking.

However, most of those require specialized hardware with infrared cameras or cameras set up as head-mounted cameras [29, 30]. Although, a few systems can estimate eye gaze through webcams. It is difficult for these systems to do it efficiently in real-life scenarios because it requires a meticulous calibration.

Baltrušaitis, Robinson, and Morency [18] indicate that, in contrast with other available options, OpenFace provides the training and test code, which is a significant advantage when reproducing experiments. OpenFace operates in real-time with all the modules concerning facial tracking analysis. Once the pupil and eye position has been detected, the CLFN model will use that information to generate five different outputs:

1. *gaze_0_x, gaze_0_y, gaze_0_z:* An eye gaze direction vector in world coordinates for eye 0 (normalized). The eye 0 is the last to the left in the image.
2. *gaze_1_x, gaze_1_y, gaze_1_z:* An eye gaze direction vector in world coordinates for eye 1 (normalized). Eye 1 is the last to the right in the image.
3. *gaze_angle_x, gaze_angle_y:* The direction of eye gaze is in radians. It is the world coordinates averaged for both eyes and converted into more user-friendly attention vectors.
4. *eye_lmk_x_0, eye_lmk_x_1,… Eye_lmk_x55, eye_lmk_y_1,… Eye_lmk_y_55*: The location of reference points of the 2D eye region in pixels (see example Fig. 3).

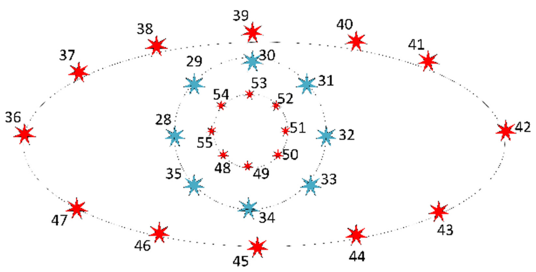

Fig. 3. Reference points to the location of the 2D eye region in pixels example [23].

5. *eye_lmk_X_0, eye_lmk_X_1,… Eye_lmk_X55, eye_lmk_Y_0,… Eye_lmk_Z_55*: The location of 3D eye region landmarks in millimeters (Fig. 4)

2.4 Additional Variables of the Robot Activity

Robot Activity

OpenFace is used to monitors in real-time the user's interaction with LOLY-MIDI EDG applications and the robotic bust. Loly collects data when it executes dialogue lines, which help evaluate its contribution to the user's interaction.

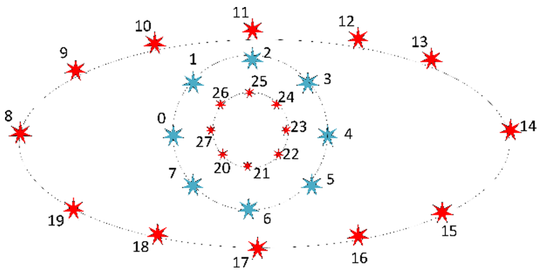

Fig. 4. Reference points location of the 3D eye region in millimeters example [23]

Four binary variables, which represent four different scenarios, are defined within the robot's programming to identify the activity it is currently on. These binary variables are based upon the line of dialogue that is being executed and whether it is talking or not: 'robot says hi,' 'robot gives instructions,' 'robot gives feedback,' 'robot is not talking.' It is worth noting that these scenarios are excluding between themselves. Binary variables will get a value depending on what line of dialogue the robot is executing to serve as conditions.

Game–Activity

A behavioral analysis of the child playing a level of the game and its activity is made, two variables are created, 'game' and 'chapter'. The first one have the name of the currently active game, for instance, Anibopi (one of the EDG of MIDI-AM series). The second one gets the name of the chapter that is active.

These variables used allows creating an analysis by game and activity. In the dashboard, the mobile application is referred to as 'game' for simplification purposes. For example, Anibopi is a game app, and one of its activities is named 'rocks. Besides making a global analysis, it is possible to filter out and analyze data for a particular game or activity satisfying the researcher's needs.

3 Results of Metric Design and Discussions

Firstly, it is necessary for data processing to determine the vectors to be analyzed concerning eye gaze and head position when the face is detected. For the eyes, the vector (gaze_angle_x, gaze_angle_y) values radians are transformed to sexagesimal degrees. For example, (3.027°, −2.718°) is obtained, understanding that they are calculated taking as a reference to the camera at (0°, 0°) [23]. In this sense, to facilitate data management, three objectives of the child's gaze are defined: robot head, Loly's bust, and tablet. The three objectives are recommended to be set as binary variables 'atent_face', 'atent_bust', 'atent_tablet', and 'atent_general'. They all take the value of 1 when the respective conditions are met and 0 if they are not.

It is considered that the child is observing the robot's head when the gaze_angle_x and gaze_angle_y angles are equal to 0°. Likewise, it is determined that the infant is looking at Loly's bust when the angle gaze_angle_x is equal to 0°, and the angle gaze_angle_y is −40°. It is established that child is looking at the tablet when: the angle gaze_angle_x is equal to 0°, and the angle gaze_angle_y is −75°. All the angles set for gaze_angle_x and gaze_angle_y can have a margin of error of 15° and 20°, respectively. Lastly, a child is considered paying attention to the interaction when variables, such as attention_face, attention_bust, or attention_tablet, are equal to 1.

If none of these conditions is true, the child is considered "not attentive". This attention metric is evaluated in real-time, and it is considered that interaction has been successful if the percentage of attention is more significant than 75% concerning the duration of the session. At least 3 h of multimodal data is collected through several sessions for each participant, including video, audio, and game preview.

3.1 Determined Correlations

Quadrant-Eye Level

The ideal position that the Loly robot camera should be placed is at the child's eye level. In this way, it is possible to capture the entire face and establish the best scenario for the use of OpenFace. This application allows us to identify the user's facial landmark, and through the collected attention metrics, it is possible to measure degrees of attention and empathy during the interaction (Fig. 5).

Fig. 5. Position of the camera concerning the children's face example

Empathy By identifying the 68 points that make up the facial landmark, the user's emotions are measured. For this purpose, specific facial points are taken as a reference to identify any reaction that can be found provoked (see Fig. 1).

Emotion Characteristics Vectors

Surprise: Increase the distance between facial points 21 and 39, the same as between facial points 22 and 42.

Interest (Neutral): Facial points 17 and 26 are at the same level or above facial points 18 and 25, respectively.

Happiness: At least one of the following conditions must be met.

1. Facial point 48 is above or at the same level as facial point 49.
2. Facial point 54 is above or at the same level as facial point 53.

Disinterest (Anger): For anger, the condition of Neutral/Interest mentioned above must be met. Also, the distance between points 62 and 66 increases considerably.

Sadness: Both conditions must be met.

1. Facial point 48 is at or below facial point 59.
2. Facial point 54 is at or below facial point 55.

Extra Variables

OpenFace can perform real-time monitoring when a child is interacting with LOLY-MIDI applications. Loly's influence in a child's interaction with the platform is measured when Loly executes lines of dialogue. It is necessary to define, within the programming of the robot metrics, four binary variables. These variables allow us to identify the robot's activity, taking into consideration the dialogue line executed and whether Loly is speaking or not.

The conditions established are as follows:

- Robot activity # 1: greeting = 1 ^ r_silence = 0
- Robot activity # 2: r_instruction = 1 ^ r_silence = 0
- Robot activity # 3: r_motivation = 1 ^ r_silence = 0
- Robot activity # 4: r_silence = 1

It is worth noting that the robot's activities are excluding. The binary variables r_greeting, r_instrution, and r_motivation must be conditioned to the dialog line's execution corresponding to the option. During activity four, to evaluate if the robot attracts the child's attention instead of playing on the tablet, three binary variables must be generated in the programming: (1) atent_loly, taking the value of 1 whenever atent_face or attention_bust is equal to 1; (2) loly_speaks, which will be 1 when r_greetings, r_instruction or r_motivation are 1; and (3) active_game, which will take the value of 1 when one of the application's games is active.

Likewise, to perform an analysis of the child's behavior by game and activity, two variables will be created, game level and chapter; the first will be the name of the active game app, for example, in Anibopi EDG. The second takes the name of the chapter that is active. Before the chapter, either of these letters will appear 'h' (game history) or 'j' (levels of the game) followed by a hyphen. These letters allow differentiating between the story part and the chapter's game. Example: 'h-rocks' or 'j-rocks'.

In the end, the outcomes presented by OpenFace generates information in 4-s intervals, as shown in the “Database” sheet (see Table 1 column 1).

Table 1. OpenFace information generated in the database Table, fragment example

Timestamp (s)	Intervention (%)	Participant	c_perpendicular	c_inc_ down	atent_face	atenc_ bust	atenc_ tablet	atenc_ general
0,00	0,00	1	1	0	1	0	0	1
4,00	0,01	1	1	0	1	0	0	1
8,00	0,01	1	1	0	0	1	0	1
12,00	0,02	1	1	0	0	1	0	1
16,00	0,03	1	1	0	0	1	0	1
20,00	0,03	1	1	0	0	1	0	1
24,00	0,04	1	1	0	0	1	0	1
28,00	0,05	1	1	0	0	0	0	1
32,00	0,05	1	0	1	0	0	1	1
36,00	0,06	1	0	1	0	0	1	1

Phonetic Analysis (Praat)

Praat makes it possible to analyze recorded audio through Loly’s microphone. Praat shows not only a graphical representation of sound through audio waves but also shows various acoustic analysis: spectrograms (a representation of the high and low frequencies available in the signal); the contour of the tone (the frequency of the periodicity); and the formation of the contour (which is the main component of the spectrogram).

Using Praat, it is intended to record and analyze harmony, intensity, frequency, and tone. Although the LOLY-MIDI application does not require the user to speak, it is possible to analyze verbal reactions to specific activities within the application while Loly speaks or not (see example Table 2).

Table 2. Percentage of Loly-MIDI attention using three users’ example

%Attention to Loly	P1	P2	P3	…	Total
Loly speaks	71,43%	80,00%	64,30%		71,43%
Loly does not speak	37,50%	25,00%	15,90%		37,50%

Metrics to measure the degree of the robotic bust support with the games

Metrics and hypotheses were determined at the end of the interaction once data is collected and processed. Also, it needs to be tried without the robot (see Table 3). A session will be considered successful if the child remains attentive to the interaction at least 75% of the time.

Table 3. Hypotheses for Loly-MIDI measures examples

Hypotheses	Phi coefficient interpretation example
Attention to Loly considerably increases when Loly speaks	0,3393
Attention to the tablet increases considerably when the game is active	0,7559
Recorded emotions generally occur when Loly speaks (loly_speaks = 1)	0,2826
Recorded emotions generally occur when the child hits/misses (r_motivation = 1)	0,4588

For example, to obtain the general confirmation of a hypothesis concerning the participants' attention, if all the participants had a 75% of attention to the interaction based on 86.6%, then the interactions are declared successful. An example of the formula to be applied is:

$$\%Atention = \frac{\#Secodswhentheatent\_general = 1}{\#Totalsecondsoftheinteraction} = \frac{130 * 4}{150 * 4} = \frac{520}{600} = 0.86 \approx 86.6\%$$

In this way, after piloting with real information, the collaboration of HCI and HRI in the attention and empathy that children with ASD can feel in the learning process can be verified.

4 Conclusions and Future Work

This paper explains the design and establishment of parameters that can capture the usability and impact of the LOLY-MIDI applications, making it possible to obtain accurate feedback reports and determine hypotheses that need to be confirmed or rejected for each case analyzed. Loly's robot autonomy linked to MIDI EDG used by children is possible due to the proposed implementation of OpenFace, which is a free facial recognition software. This method allows the LOLY-MIDI platform to carry on with the intervention without the need for a present real-time observer that evaluates attention and empathy levels. Data collected in real-time by OpenFace can be evaluated using pre-established conditions that dictate an intervention's success or failure.

The lessons learned through the pilot tests were carried out but could not be processed using OpenFace. However, they still contributed to the reformulation of the initial hypotheses regarding the empathy and attention of a child, along with interviews with specialized psychologists in children with ASD that helped determine the limitations of LOLY-MIDI applications, turning this work into a solid baseline on which to continue researching. Furthermore, this work can contribute to the development of metrics in different robotic platforms used to perform HRI since this work could provide a guideline to measure the degree of attention that a person can have because many times, the development of robotic platforms is truncated. After all, the user does not understand the robot's instructions due to lack of attention.

4.1 Future Work

After establishing metrics and the HRG platform's ways designed and discussed, the next stage of this research is developing a computer module; this module needs to be programmed to computerize the designed metrics to automatically present results for each evaluation. Also, the implementation of OpenFace using Loly's hardware must be exhaustively tested and calibrated to get reliable data output and make an accurate data analysis considering the pre-established conditions for all the binary variables to help us evaluate every aspect of the interaction. A thorough evaluation of the proposed LOLY-MIDI HRG platform is required. For evaluation purposes, data collection should need a representative sample size to evaluate the determined hypotheses. It is also recommended that if there is the possibility of modifying the platform to generate interaction between the child and the robot Loly, the impact will be made and measured by Praat phonetic analysis because although it is not currently part of the proposal in the short term, it is considered that it can become instrumental in learning children.

References

1. Gros, B.S.: Digital games in education: the design of games-based learning environments. J. Res. Technol. Educ. **40**(1), 23–38 (2007)
2. Plass, J.L., Homer, B.D., Kinzer, C.K.: Foundations of game-based learning. Educ. Psychol. **50**(4), 258–283 (2015)
3. Feil-Seifer, D., Mataric, M.J.: Defining socially assistive robotics. In: 9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005, pp. 465–468 (2005)
4. Paillacho Chiluiza, D.F., Solorzano Alcivar, N.I., Paillacho Corredores, J.S.: LOLY 1.0: a proposed human-robot-game platform architecture for the engagement of children with autism in the learning process. In: Botto-Tobar, M., Zamora, W., Larrea Plúa, J., Bazurto Roldan, J., Santamaría Philco, A. (eds.) ICCIS 2020. AISC, vol. 1273, pp. 225–238. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-59194-6_19
5. Volkmar, F.R., Paul, R., Rogers, S.J., Pelphrey, K.A.: Handbook of Autism and Pervasive Developmental Disorders, Diagnosis, Development, and Brain Mechanisms. John Wiley & Sons, 2014 (2014)
6. Rice, K., Moriuchi, J.M., Jones, W., Klin, A.: Parsing heterogeneity in autism spectrum disorders: visual scanning of dynamic social scenes in school-aged children. J. Am. Acad. Child Adolesc. Psychiatry **51**(3), 238–248 (2012)
7. Tsatsanis, K.D.: Heterogeneity in learning style in Asperger syndrome and high-functioning autism. Topics Lang. Disord. **24**(4), 260–270 (2004)
8. American-Psychiatric-Association: Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub, 2013 (2013)
9. Sigman, M., Mundy, P., Sherman, T., Ungerer, J.: Social interactions of autistic, mentally retarded and normal children and their caregivers. J. Child Psychol. Psychiatry **27**(5), 647–656 (1986)
10. Christensen, D.L., et al.: Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2012. MMWR Surveill. Summ. **65**(13), 1 (2018)
11. Ministerio de Salud Pública del Ecuador: Trastornos del Espectro Autista en niños y adolescentes: Detección, diagnóstico, tratamiento, rehabilitación y seguimiento. Guía de Práctica Clínica (2017)

12. Kozima, H., Nakagawa, C., Yasuda, Y.: Interactive robots for communication-care: a case-study in autism therapy. In: ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, pp. 341–346 (2005)
13. Contreras, R.S., Serra, A., Terrón, J.L.: Games and ADHD-ADD: a systematic mapping study. Acta Ludologica **2**(2), 4–26 (2019)
14. Yamakami, T.: From user experience to social experience: A new perspective for mobile social game design. In: 2012 9th International Conference on Ubiquitous Intelligence and Computing and 9th International Conference on Autonomic and Trusted Computing, pp. 792–796 (2012)
15. Solorzano, N.I., Carrera, D.A., Sornoza, L.I., Mendoza, M.: Developing a dashboard for monitoring usability of educational games apps for children. In: Proceedings of the 2019 2nd International Conference on Computers in Management and Business, pp. 70–75 (2019)
16. Solorzano, N.I., Sornoza, L.I., Carrera, D.A.: Adoption of children's educational video games monitored with dashboards in the cloud. Rev. Iberica Sist. Technol. Inf. **2019**(19), 146–160 (2019)
17. Jain, S., Thiagarajan, B., Shi, Z., Clabaugh, C., Matarić, M.J.: Modeling engagement in long-term, in-home socially assistive robot interventions for children with autism spectrum disorders. Sci. Robot. **5**(39) (2020)
18. Baltrušaitis, T., Robinson, P., Morency, L.-P.: Openface: an open-source facial behavior analysis toolkit. In: IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1–10. IEEE (2016)
19. Solorzano, N.I., Sanzogni, L., Houghton, L.: A pluralistic methodology for a refined selection of drivers influencing information system adoption in public organizations: the case for Ecuador (2016)
20. Creswell, J.W., Creswell, J.D.: Research design: qualitative, quantitative, and mixed methods approaches. Sage publications, 2017, Third Edition edn. (2017)
21. Baltrusaitis, T., Robinson, P., Morency, L.-P.: Constrained local neural fields for robust facial landmark detection in the wild. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 354–361 (2013)
22. Cristinacce, D., Cootes, T.F.: Feature detection and tracking with constrained local models. In: Bmvc, Vol. 1, No. 2, p. 3 (2006)
23. Baltrusaitis, T.: Output Format OpenFace. In: Editor (Ed.)ˆ(Eds.): 'Book Output Format OpenFace' (edn.), pp. FaceLandmarkImg and FeatureExtraction (2019)
24. Yu, Y., Mora, K.A.F., Odobez, J.-M.: Robust and accurate 3d head pose estimation through 3dmm and online head model reconstruction. In: 12th IEEE international conference on automatic face & gesture recognition (fg 2017), pp. 711–718 (2017)
25. Fanelli, G., Gall, J., Van Gool, L.: Real-time head pose estimation with random regression forests. In: CVPR 2011, pp. 617–624. IEEE (2011)
26. Asthana, A., Zafeiriou, S., Cheng, S., Pantic, M.: Robust discriminative response map fitting with constrained local models. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3444–3451 (2013)
27. Asthana, A., Zafeiriou, S., Cheng, S., Pantic, M.: Incremental face alignment in the wild. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1859–1866 (2014)
28. https://medium.com/insights-on-virtual-reality/what-you-should-know-about-head-trackers-7e2289578a22. Accessed 11 Oct 2020

29. Świrski, L., Bulling, A., Dodgson, N.: Robust real-time pupil tracking in highly off-axis images. In: Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 173–176 (2012)
30. Lidegaard, M., Hansen, D.W., Krüger, N.: Head-mounted device for point-of-gaze estimation in three dimensions. In: Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 83–86 (2014)