A proposal of a virtual robotic assistant and a rule-based expert system to carry out therapeutic exercises with children with Dyslalia

V. Robles-Bykbaev, V. Velásquez-Angamarca, K. Mosquera-Cordero,
D. Calle-López, Y. Robles-Bykbaev, E. Pinos-Vélez
GI-IATa, Cátedra UNESCO Tecnologías de apoyo
para la Inclusión Educativa, Universidad Politécnica Salesiana
Cuenca, Ecuador
Email: vrobles@ups.edu.ec, {vvelasqueza,kmosquerac}@est.ups.edu.ec,
{dcallel,zrobles,epinos}@ups.edu.ec

A. León-Pesántez
Escuela de Educación Especial
Universidad del Azuay
Cuenca, Ecuador
Email: aleon@uazuay.edu.ec

***Abstract*—The Dyslalia is considered one of the most common communication disorders that present children with and without disabilities. The National Institute on Deafness and Other Communication Disorders (NIDCD) claims that 5 percent of U.S. children ages 3-17 have a speech disorder that lasted for a week or longer during the past 12 months. This situation is very similar in other countries of Europe, Africa, and Latin America. For these reasons, in this paper, we present a proposal of a virtual robotic assistant aimed at providing therapy exercises for children with Dyslalia. The virtual assistant uses a rule-based expert system that determines the best alternatives of exercises that will be carried out with a child in the phonetic area. With the aim of developing the physical version of the virtual assistant, we performed a study with 33 students enrolled in the last year of the career of initial education, early stimulation, and precocious intervention (in Cuenca, Ecuador). The participants validated our proposal through a survey, and the achieved results are encouraging.**

***Index Terms*—Virtual Robotic Assistant, Dyslalia, Expert System, Speech Therapy, Rules-Based Reasoning.**

I. Introduction

The language is considered as a mainstay for the cognitive, social and affective development in children. Therefore, the language allows children exchange information and express feelings, thoughts, and emotions to break through a world of knowledge and opportunities [1], [2].

During the early years, a child acquires and develops complex components of language in areas such as the phonetic-phonological, semantic, morphosyntactic and pragmatic. This process - in subjects without disabilities - is carried out naturally given that children continuously receive stimuli from the environment [1], [2].

The phonetic-phonological component - the objective of this research - is focused on the utterances production in a given language, this is, the phonemes articulation. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM5), the phonological production describes the appropriate articulation of phonemes (i.e., individual sounds) that are combined to create spoken words. Phonological production requires both the phonological knowledge of speech sounds and the ability to coordinate the movements of the organs of the joint (jaw, tongue, and lips) with breathing and vocalization [3].

According to the research carried out by [4] in 200 of Mexican children, the phonetic acquisition varies according to the age. In this line, a child can adequately pronounce the phonemes $[m, ch, n, k, t, y, p, l, f, d, j]$ at three years, $[b, g, r, bl, pl, s]$ at four, $[kl, br, fl, kr, gr]$ at five, and $[rr, pr, gl, fr, tr]$ at six. Therefore, at the age of four, a child has acquired the most of phonemes, with some exceptions, and especially the /rr/ for Spanish (that is developed until 5 or 6 years).

During his/her normal development, a child gradually acquires different sets of phonemes. However, if a child can not adequately articulate the phonemes that correspond to his/her age, a Speech-Language Therapist (SLT) diagnosis this problem as the communication disorder named Dyslalia. The literature defines the Dyslalia as a "disorder caused by organic or functional impairments of the speech peripheral organs, which consists of the impossibility to correctly utter one or several sounds for the emission of a certain phoneme." [5], [6]. It is essential mentioning that this disorder can appear in both spontaneous and repetitive speech.

In the same line, [7] points that Dyslalia is a disorder in the phonemes' articulation, either by the absence or alteration of specific sounds or by the substitution of these by others in an inappropriate way. It is an inability to pronounce certain phonemes or groups of phonemes correctly.

On the other hand, [8] claims that the Dyslalia could have a prevalence between 5 and 10% in the infantile population. Otherwise, [9] places Dyslalia in the three most frequent infantile language disorders (joint with the dysphonia and language delay).

Similarly, [10] points out that the Dyslalia is the problem that has the highest prevalence as well as it is the disorder for

978-1-5386-6657-9/18/$31.00 ©2018 IEEE

which the most of parents and relatives request support. This situation commonly occurs due to the relationship that exists between Dyslalia and some difficulties related to learning to read.

For these reasons, it is fundamental to provide an early intervention on those children that present Dyslalia. With this intervention, patients will be able to correctly pronounce the sounds of language and avoid the involution and intensification of the difficulties presented [11].

Finally, it is essential to remark that [12] indicates this disorder can persist in adulthood, implying in some cases problems to get jobs, performing work activities or having difficulties in social adjustment.

In the light of the above, in this paper, we present a proposal of a robotic virtual assistant to support the educational and rehabilitation activities of children with Dyslalia. In this first stage, we have determined the perception of 33 students regarding the pertinence and usefulness of the virtual assistant as well as to its 3D printed version.

II. Related work

The Dyslalia is a speech disorder for which nowadays exist several proposals. In this line, in [13] the authors present an interactive toolset with the objective of combining biofeedback, self-monitoring, and reward mechanisms in an entertainment environment focused on the speech therapy. The toolset relies on visual biofeedback with the aim of motivating patients, tracking orofacial motion, and evaluating the patients' speech productions.

Similarly, [14] present a serious game designed to provide speech therapy for Egyptian speaking children with Dyslalia. To this aim, the authors developed a speech engine to accurately and automatically detect the problem in the child pronunciation. The engine was trained with a speech resource corpus based on the Modern Standard Arabic and adapted to Egyptian dialect. The authors report accuracy levels of 93.4% and 95.6% to recognize the phonemes /s/ and /r/, respectively.

In the area of the artificial intelligence, [15] developed an expert system based on fuzzy logic to optimize Computer-Based Speech Therapy systems. The primary objectives of this system are the following: a) provide personalized therapy (following childs problems level, context, and his/her possibilities), b) give support to therapists (offer some suggestions regarding what type of exercises are better for a specific moment and a specific child), and c) self-teaching (change the knowledge base when the system's suggestion differ from the viewpoint of therapists). The rules of the expert system are defined using the following input variables: the defect of speech, child's age, and family involving. The output variable is established as the number of weekly therapy sessions. To validate their system, the authors worked with 20 children with ages between five and six years. The authors have designed two groups of 10 children, and for the first one, the system established the number of working sessions whereas for the second group this task was carried out by a therapist. The results reported in this research show that both groups progressed, and both groups reached the same performance [15].

Other remarkable contributions related with Dyslalia and in general, speech disorders, are described below:

- A proposal for a virtual world to supports therapy activities for Dyslalia is presented by [16]. This system focuses on the articulatory reeducation with phonemes and symphons. The children can interact with an avatar, which can remain in the virtual world along with all other users after logging.
- The research conducted by [17] explores the possibility of performing multimodal emotion recognition with the aim of integrating this technique in computer-based speech therapy systems. To accomplish this objective, the authors use a multimodal emotion recognition framework that relies on video, audio, and physiological channels (breathing, temperature, skin conductance, and heart rate/blood volume pulse).
- The telerehabilitation of children with Dyslalia was analyzed by [18]. In this research, the authors discuss the design of educational videos for telerehabilitation. To create the videos the authors used the following principles of multimedia learning:
  - Coherence: design elements like logo graphic and intro/closure music were not added throughout the video.
  - Signaling: in order to highlight essential words or movements, visual signaling was provided in the form of arrows, accented text or other visual clues.
  - Redundancy: exclude redundant captions from narrated animation.
- A system for automatic assessment of articulatory deviations in pathological speech (Serbian) was developed by [19]. To this aim, the authors implemented a module that uses Artificial Neural Networks (ANN). This module which corresponds to the Global Articulation Test (GAT) uses the phoneme as the input and gives the quantitative assessment of deviation as a result. To train the ANN, the authors created a speech database with the support of children and adults. Each subject pronounced 30 times isolated words from the stimuli list of the GAT. After this, the authors implemented an automatic segmentation algorithm based on the DTW (Dynamic Time Warping) and kNN (k nearest neighbors) algorithms. This algorithm segmented the speech database in words to feed the ANN. The authors report 96% of precision for the recognition task of articulatory-acoustic deviations.

III. Proposed approach

In this section, we will describe the most relevant aspects of our proposal, paying special attention to the virtual robotic assistant, and the expert system based on rules. Below we describe the most relevant modules as well as its functionalities (Figure 1):

- To interact with the virtual robotic assistant, the system provides children a **virtual access interface** that is suitable for several platforms such as **mobile devices**

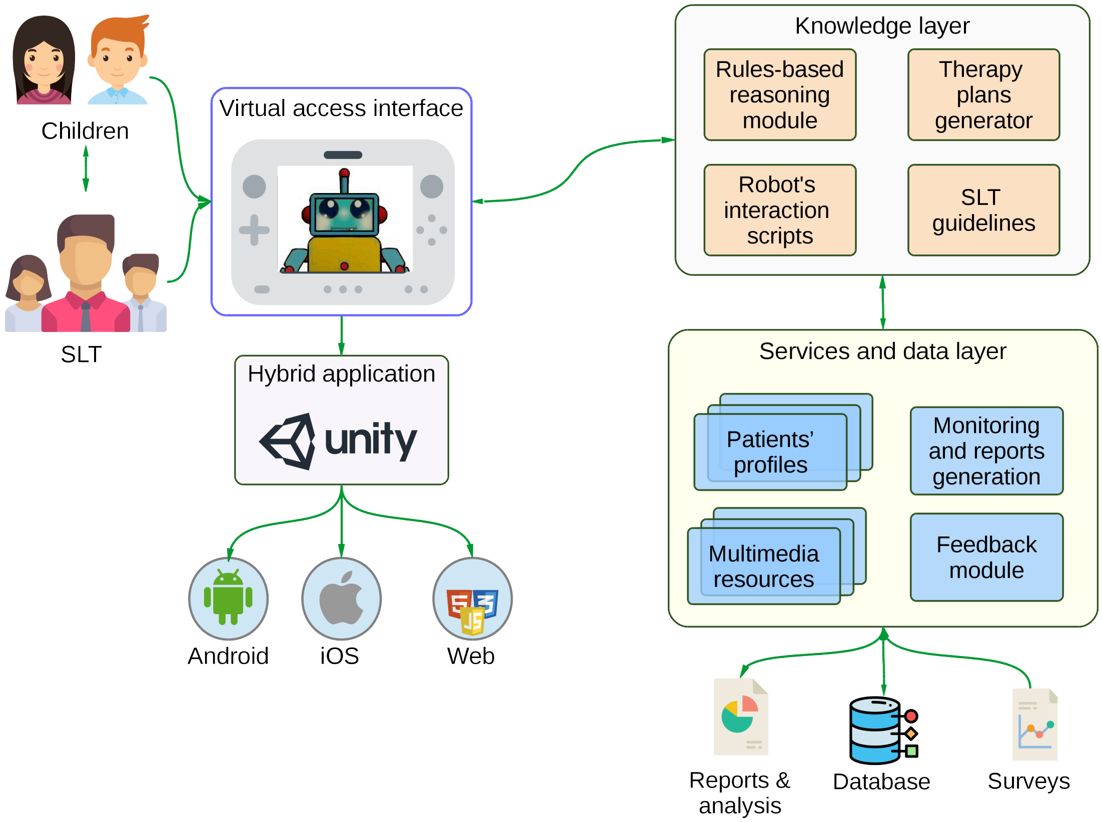

Fig. 1. The general architecture of the proposed system and its main components and modules.

(Android®, and iOS®) and **desktop environments**. This interface was developed using Unity® which is why it can be compiled and deployed in the platforms previously mentioned.

- With the aim of providing support to SLTs, the system includes a **knowledge layer** which is responsible for suggesting intervention activities. This layer achieves this goal by a **rule-based reasoning module** that represents through facts and rules the intervention guidelines. The reasoning module is used by **therapy plans generator** to assemble an intervention plan with suggested exercises according to child's profile. In the same way, the plans generator uses the module of **robot's interaction scripts** to establish how will be the interaction with the children.
- **The services /data layer** contains several functionalities to manage the information. In this layer are stored the **patients' profiles** which includes the following data: child's personal information (names, age, gender, etc.), and medical profile (SLT initial evaluation and diagnosis). Other essential elements of this layer are the service for **monitoring the patient's progress** and generate reports and a **feedback module** that will be implemented in the future to collect an evaluation (made by SLTs) of the therapy plans created by the system.

A. Therapeutic activities sequence logics

With the objective of carrying out the therapy with children, the virtual robot uses the sequence depicted in Figure 2.

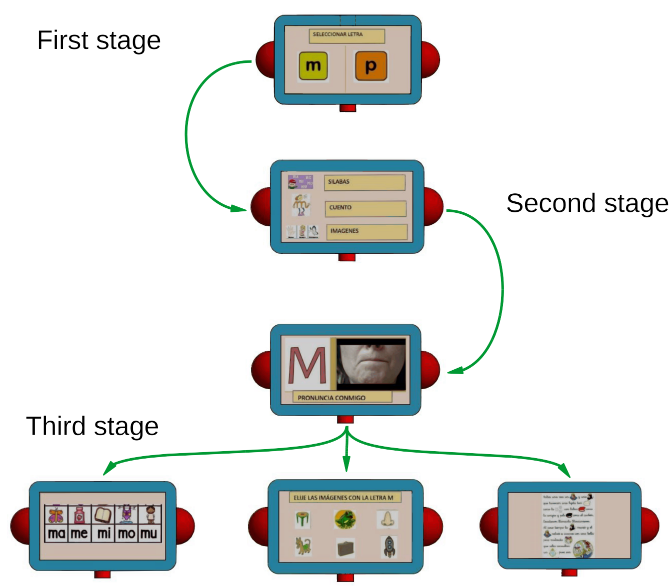

Fig. 2. An example of the sequence of therapeutic activities presented by the virtual robotic assistant.

As it can be seen, the sequence commonly has three stages.

In the first one, the therapist must choose with which phonetic unit will be done the exercises. After that, the therapist can select to work with one of the following categories that contain the phonetic unit selected: syllables, a story, or pictograms. In the third stage, the therapist will select the specific activity to carry out with children. For example, if the therapist chooses syllables, it would be possible to work with utterances production. It is important to mention these activities will show in the robots face, and when the child responds, the system will give him/her an appropriate stimulus according to whether the answer was correct or not).

B. Rules-based reasoning module for therapy plans design

An expert system can help us solve problems related to the knowledge domain with which it is assembled. The primary objective of the proposed expert system is to recommend exercises for the treatment of specific conditions related to Dyslalia (Figure 3).

Fig. 3. A partial screen capture of the rules used (top image) to generate therapy plans (bottom image) for children with Dyslalia.

Based on the exercises proposed by experts in the area, we proceeded to make an expert system using the Prolog programming language. The rules and the knowledge base have been defined based on the cases given in our experimentation. The knowledge base comprises mainly children around five years old with different types of diagnoses related to Dyslalias.

The results can be used to provide support to professionals so that, based on the results obtained within the evaluated areas, a diagnosis can be obtained with several recommended exercises to solve the problems related to the detected Dyslalias (rhotacism, lambdacism, sigmatism, etc.). Having that:

- $D \rightarrow$ diagnosis
- $AE \rightarrow$ evaluated area
- In the indicated algorithm (Figure 3), the pairs D_p, AE_q represent the evaluated area AE_q corresponding to the D_p diagnosis.
- For each of the areas evaluated, if the expert system receives a positive response, proceed with the following AE.
- If the assertion is *false*, proceed to AE_1 of the next diagnostic group $D(p+1)$
- When the cases AE_1 to AE_n are met, the current D diagnosis is considered correct and the related exercises are recommended.

IV. EXPERIMENT AND RESULTS

With the objective of determining the real feasibility of including our proposal as a support tool to carry out therapy activities for Dyslalia treatment, we conducted a pilot experiment with 33 students assisting to last semester of the career of initial education, early stimulation, and precocious intervention (University of Azuay, Cuenca - Ecuador).

Below we present some of the most relevant characteristics of the sample of students that participated in the experiment:

- All students had experience working with children that presented different kinds of speech disorders (including Dyslalia).
- All the volunteers are women with ages between 21 and 26 years.
- The participants have experience working with commercial informatics tools that commonly are used during the therapy sessions.

To achieve our objective, we presented the system to the volunteers, and after that, we asked them to respond to a survey that has the following characteristics:

- Is organized in two sections, one to collect the personal data and professional profile of the participants, and a second that consists of a set of questions about the virtual robotic assistant.
- The set of questions focuses on the following aspects: appearance and shape of the virtual robot, the colors of the avatar, the sequence in which the therapeutic activities are executed (if it is appropriate or not), the robot's expressions, the desired features of a 3D printed version of the robot, among many others.
- Each question was designed using the Likert scale (5 = strongly agree, 4 = agree, 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree) [20].

As a previous step to apply the survey, we validated it. The reliability of this survey was determined by the Cronbach's Alpha test [21], whose index was 0.83. The value obtained allows us to indicate that there is internal coherence between the items, that is, the survey is reliable.

Once we performed this examination, the next step was carry out the descriptive analysis, for which we used the R Language and the free software "RStudio," with versions 3.5.0 and 1.0.136, respectively.

As it can be seen in Figure 4, all the surveyed volunteers have "positive" and "very positive" perception about the usefulness of the virtual robotic assistant as a support tool in the treatment of children with Dyslalia.

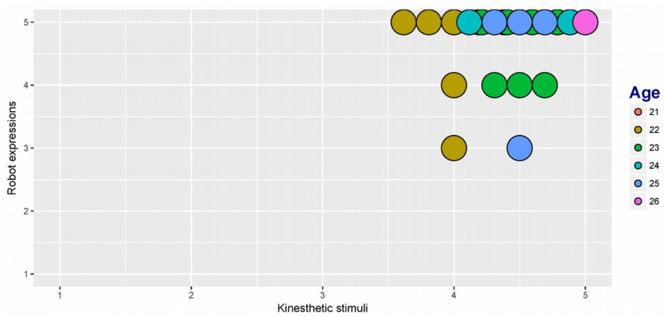

Fig. 4. The perception levels of the experiment's volunteers regarding the virtual robotic assistant.

On the other hand, in Figure 5 we can see the volunteers criterion about the appropriateness of the therapeutic activities sequences. The participants of the experiment consider the robot as a useful support tool for therapies. In the same way, the volunteers consider that robot's appearance is "very appropiate" to work with children.

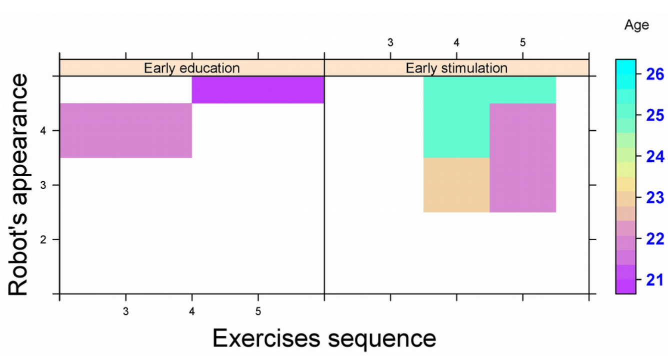

Fig. 5. The perception levels of the experiment's volunteers regarding the appropriateness of the activities' sequences and virtual robotic assistant usefulness.

In the light of the above, we can asseverate that our approach has an adequate level of acceptation by the groups of future professionals in charge of the educational and rehabilitation processes of children with speech disorders.

V. Conclusion

In this paper, we presented a proposal of a virtual robotic assistant to support the rehabilitation and educational activities of children diagnosed with Dyslalia. Our proposal has the novelty that combines an avatar (robot) with a rules-based reasoning system. This last allows us changing the avatar for any other character that could more appropriate to work with children with disabilities (autism spectrum disorder, intellectual disability, etc.).

Another important aspect related to our proposal is that the expert system can be extended to cover other speech-language disorders through the incorporation of new rules and facts. Finally, it is relevant to remark that the future professionals that will work with children with Dyslalia and other disorders consider our system as a useful support tool.

As lines of future work we propose the following ones:

- To develop a 3D printed version of the robotic assistant.
- To include serious games in the activities provided by the virtual robotic assistant.

Acknowledgment

This work has been funded by the "Sistemas Inteligentes de Soporte a la Educación (v5)" research project, the Cátedra UNESCO "Tecnologías de apoyo para la Inclusión Educativa" initiative, and the Research Group on Artificial Intelligence and Assistive Technologies (GI-IATa) of the Universidad Politécnica Salesiana, Campus Cuenca.

References

[1] C. S. Tamis-LeMonda, J. D. Shannon, N. J. Cabrera, and M. E. Lamb, "Fathers and mothers at play with their 2-and 3-year-olds: Contributions to language and cognitive development," *Child development*, vol. 75, no. 6, pp. 1806–1820, 2004.

[2] T. E. Moore, *Cognitive Development and Acquisition of Language*. Elsevier, 2014.

[3] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[4] M. Melgar de González, *Cómo Detectar Al Niño Con Problemas Del Habla*. Editorial Trillas, 2012.

[5] M. L. Alessandri, "Trastornos del Lenguaje, Detección y Tratamiento en el Aula," *Buenos Aires: Landeira Ediciones SA*, 2011.

[6] M. M. Cruz-Cunha, *Handbook of Research on Serious Games as Educational, Business and Research Tools*. IGI Global, 2012.

[7] P. Pascual García, *La dislalia: naturaleza, diagnóstico y rehabilitación*. CEPE. Ciencias de la Educación Preescolar y Especial, 2012.

[8] P. Conde-Guzón, P. Quirós-Expósito, M. J. Conde-Guzón, and M. T. Bartolomé-Albistegui, "Perfil neuropsicológico de niños con dislalias: alteraciones mnésicas y atencionales," *anales de psicología*, vol. 30, no. 3, pp. 1105–1114, 2014.

[9] M. Escabias Gutiérrez, "Los trastornos del lenguaje en los niños," *Revista Digital Enfoques Educativos*, vol. 106, 2008.

[10] M. García Mateos, "Análisis de la demanda asistencial en trastornos del lenguaje, habla y comunicación. Un estudio epidemiológico realizado desde la unidad asistencial de logopedia de la UPSA," Ph.D. dissertation, Departamento de Psicologa Evolutiva y de la Educación. Universidad de Salamanca, 10 2010.

[11] E. Włodarczyk, A. Szkiełkowska, H. Skarżyński, and A. Piłka, "Assessment of the efficiency of the auditory training in children with dyslalia and auditory processing disorders," *Otolaryngologia polska= The Polish otolaryngology*, vol. 65, no. 5, pp. 339–344, 2011.

[12] B. Schonhaut and L. Maggiolo, "Alteraciones del lenguaje en la población preescolar," *Otorrinolaringologia. Suplemento de Salud (i) Ciencia*, pp. 1667–8982.

[13] A. Grossinho, S. Cavaco, and J. Magalhaes, "An interactive toolset for speech therapy," in *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*. ACM, 2014, p. 36.

[14] R. Elhady, M. Elmahdy, I. Hamed, and S. Abdennadher, “A game with a purpose for automatic detection of childrens speech disabilities using limited speech resources,” in *International Conference on Arabic Language Processing*. Springer, 2017, pp. 79–89.

[15] O. A. Schipor, S. G. Pentiuc, and M. D. Schipor, “Improving computer based speech therapy using a fuzzy expert system,” *Computing and Informatics*, vol. 29, no. 2, pp. 303–318, 2012.

[16] E. Q. Rivas and E. S. Molina, “A proposal for a virtual world that supports theraphy of dyslalia,” in *Proceedings of the 6th Euro American Conference on Telematics and Information Systems*. ACM, 2012, pp. 371–374.

[17] O.-A. Schipor, S.-G. Pentiuc, and M.-D. Schipor, “Towards a multimodal emotion recognition framework to be integrated in a computer based speech therapy system,” in *Speech Technology and Human-Computer Dialogue (SpeD), 2011 6th Conference on*. IEEE, 2011, pp. 1–6.

[18] D. P. Vukovac, “Employing educational video in speech-language telerehabilitation with children: insights from interviews with parents,” in *8th International Conference on Education and New Learning Technologies*, 2016.

[19] R. B. Bilibajkić, “Recognition of articulatory-acoustic deviations in pathological speech,” in *Telecommunications Forum (TELFOR), 2016 24th*. IEEE, 2016, pp. 1–8.

[20] J. A. Gliem and R. R. Gliem, “Calculating, interpreting, and reporting cronbachs alpha reliability coefficient for likert-type scales.” Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, 2003.

[21] J. R. A. Santos, “Cronbachs alpha: A tool for assessing the reliability of scales,” *Journal of extension*, vol. 37, no. 2, pp. 1–5, 1999.