Psicología Educativa
ISSN: 1135-755X
ISSN: 2174-0526
Colegio Oficial de la Psicología de Madrid

González, Rocío; Zalazar-Jaime, Mauricio F.; Medrano, Leonardo A.

Assessing Complex Emotion Recognition in Children.
A Preliminary Study of the CAM-C: Argentine Version

Psicología Educativa, vol. 30, no. 1, 2024, pp. 19-28
Colegio Oficial de la Psicología de Madrid

DOI: https://doi.org/10.5093/psed2024a4

Available in: https://www.redalyc.org/articulo.oa?id=613777779003

Psicología Educativa (2024) 30(1) 19-28

Psicología Educativa

https://journals.copmadrid.org/psed

Assessing Complex Emotion Recognition in Children. A Preliminary Study of the CAM-C: Argentine Version

Rocío González[1], Mauricio F. Zalazar-Jaime[2], and Leonardo A. Medrano[1,3]

[1]Universidad Siglo 21, Argentina; [2]Universidad Nacional de Córdoba, Argentina; [3]Pontificia Universidad Católica Madre y Maestra, República Dominicana

ARTICLE INFO

Article history:
Received 31 March 2023
Accepted 31 October 2023

Keywords:
Facial emotion recognition
CAM-C test
Psychometric properties
Spanish-speaking population

ABSTRACT

Facial emotion recognition is one of the psychological processes of social cognition that begins during the first year of life, though the accuracy and speed of emotion recognition improves throughout childhood. The objective of this study was to carry out a preliminary study for the adaptation and validation of the CAM-C FACE test in Argentine children from 9 to 14 years old, by measuring hit rates and reaction times. The results of this study show that the unidimensional model is more appropriate when assessing the speed of performance (reaction times), with a satisfactory reliability (ρ = .950). Results also indicated that girls presented more correct answers compared to boys, while boys had longer reaction times. In addition, the group of children from 12 to 14 years old presented more correct answers compared to the group from 9 to 11 years old, while no differences were observed between groups in terms of reaction times.

Evaluación del reconocimiento de emociones complejas en niños. Un estudio preliminar del CAM-C: versión argentina

Palabras clave:
Reconocimiento facial de emociones
Test CAM-C
Propiedades psicométricas
Población de habla española

RESUMEN

El reconocimiento facial de emociones es uno de los procesos psicológicos de la cognición social que comienza durante el primer año de vida, aunque la precisión y la velocidad de reconocimiento emocional mejora a lo largo de la infancia. El objetivo de esta investigación fue realizar un estudio preliminar de la adaptación y validación del test CAM-C FACE en niños argentinos de 9 a 14 años de edad, evaluando las respuestas correctas y los tiempos de reacción. Los resultados mostraron que el modelo unidimensional es el más apropiado cuando se mide la velocidad de ejecución (tiempos de reacción), con una confiabilidad satisfactoria (ρ = .950). Los resultados también indicaron que las niñas presentan más respuestas correctas que los niños, mientras que estos tienen tiempos de reacción más largos. Asimismo, el grupo de niños de 12 a 14 años presentan más respuestas correctas que el de 9 a 11 años, mientras que no se observan diferencias entre grupos de edad en el tiempo de reacción.

Social cognition is defined as a set of (conscious and non-conscious) psychological processes that underlie social interactions. It includes the mental operations that are involved in the perception, interpretation, and generation of responses to the intentions, dispositions, and behaviors of others. This allows us to understand, act, and benefit from the interpersonal world (Kennedy & Adolphs, 2012; Vatandoust & Hasanzadeh, 2018). The neural network that underlies these abilities was first described by Brothers and Ring (1992) as "the social brain". The medial, inferior frontal, and superior temporal cortices, along with the amygdala, form a network of brain regions that implement computations relevant to social processes. Perceptual inputs to these social computations may arise in part from regions in the fusiform gyrus and from the adjacent inferior occipital gyrus that activate in response to faces (Golan et al., 2006).

Since emotions use non-verbal signals as the main vehicle for their expression, one of the most basic and widely studied social cognition processes is the recognition of emotions through non-verbal communication (Leiva, 2017; Lieberman 2010). Although the recognition of the emotions and mental states in others depends on the ability to integrate multimodal information in context (facial expression, vocal intonation, body language, contextual information), most studies on emotion recognition have focused specifically on facial expression due to the centrality that it has in emotional expression (Fridenson-Hayo et al., 2016; Ko, 2018).

Cite this article as: González, R., Zalazar-Jaime, M. F., & Medrano, L. A. (2024). Assessing complex emotion recognition in children. A preliminary study of the CAM-C: Argentine version. *Psicología Educativa, 30*(1), 19-28. https://doi.org/10.5093/psed2024a4

Funding: This research work was supported in part by grants awarded to the first author from the "Consejo Nacional de Investigaciones Científicas y Técnicas" (CONICET), Argentina. The funding sponsors had no involvement in study design, collection, analysis, or interpretation of data, writing the manuscript, nor in the decision to submit the manuscript for publication. This study was not preregistered. Data and study materials are not available to other researchers. Correspondence: mgrociogonzalez@gmail.com (R. González).

ISSN: 1135-755X/

The adaptive value of facial expressions of basic emotions is that they reliably show the emotional state of a person, along with their behavioral tendency. This directly influences the establishment and regulation of social interactions (Damasio, 2010; Izard, 1977; Ko, 2018). Research has mainly focused on the recognition of six emotions that are considered "basic" (happiness, sadness, fear, anger, surprise, and disgust). These "basic emotions" are cross-culturally expressed and recognized and, to some extent, are neurologically distinct (Adolphs, 2003; Ekman & Cordaro, 2011; Wilson-Mendenhall et al., 2013).

In addition to basic emotions, there are complex emotions, which are intended to coordinate social interactions, regulate relationships, and maintain group cohesion, which is also essential for survival (Keltner & Haidt, 1999). According to LeDoux (2000), complex emotions arise from the combination of basic emotions, so the recognition of these emotional states requires a cognitive elaboration of the social context. Complex emotions take place from social interaction, involve attributing cognitive states and emotions to others, and are more context and culture dependent (Golan et al., 2015).

Basic emotions would be recognized from the earliest years of life, though typically developing children begin to recognize and verbally label complex emotions, such as shame, pride, and jealousy, by the age of 7 years (Iglesias et al., 1989; Walle et al., 2020). Although the ability to discriminate emotions begins during the first year of life, the accuracy and speed of emotion recognition improves throughout childhood, because the abilities to recognize emotions and mental states continue to develop in adolescence and adulthood (Golan et al., 2015). De Sonneville et al. (2002) showed that in the age range of 7-10 years accuracy of facial processing hardly increased, while speed did substantially increase with age. Adults, however, were substantially more accurate and faster than children. They conclude that speed is a more sensitive measure when children get older and that speed of performance, in addition to accuracy, might be successfully used in the assessment of clinical deficits.

This is the reason why it is important to record not only hit rates (i.e., the percentage of right answers), but also reaction times of facial emotion recognition, since this measure could provide a better differentiation between participants and between stimuli (Kosonogov & Titova, 2019). For example, in their review of 29 studies of schizophrenia, Edwards et al. (2002) found only six studies that measured reaction times of facial emotion recognition, showing that patient deficits could be identified, among other things, by measuring reaction times. In the Argentine context, a study carried out in a child and adolescent population found that women perform better not only in the accuracy of basic emotional recognition, but also in processing speed (Morales et al., 2017).

Although the progress of neuropsychological research on non-verbal communication and emotional processing is remarkable, little progress has been made in the development of standardized measures for the study of individual differences in the recognition of complex facial expressions (Suzuki et al., 2006). On the other hand, emotional recognition assessment tests have mainly used prototypical static facial expressions images, mostly taken from standardized tests such as the Pictures of Facial Affect by Ekman and Friesen (1976) or the Japanese and Caucasian Facial Expressions of Emotion by Matsumoto and Ekman (1988) (Golan et al., 2015). However, the use of static stimuli presents an ecological limitation, since these are very different from the way emotional gestures are presented in everyday life (Kosonogov & Titova, 2019). In a review by Krumhuber et al. (2013) it was observed that dynamic information improves coherence in the identification of emotion and helps differentiate between genuine and fake expressions. Dynamic properties of facial stimuli have been shown to influence the processing of emotional information (Recio et al., 2013), to such an extent that patients with brain injuries obtain poorer recognition performance when the stimuli are static, while the addition of dynamic information increases the number of correct answers in recognition (Adolphs et al., 2003; McDonald & Saunders, 2005; Zupan & Neumann, 2016). Therefore, evaluation using static stimuli would overestimate the deficits in patients by not including dynamic information present in daily life (Leiva, 2017).

In this Field Golan et al. (2015) designed a battery called the Cambridge Mindreading Face-Voice Battery for Children (CAM-C). This battery assesses the recognition of nine complex emotions in children through facial expressions. An advantage compared to other tests is that it not only uses dynamic stimuli, but that coloured full-face video clips are represented by adults and children. Although reaction times allow for greater accuracy, to our knowledge it has never been examined for this test. On the other hand, no psychometric validation studies have been carried out in a Spanish-speaking population. In addition to this, no studies have been found that have analyzed the factorial structure of the test, thus the mono-dimensional structure of the test has never been formally investigated by factor analysis (Barceló-Martínez et al., 2018; Golan et al., 2015; Rodgers et al., 2021).

Accordingly, the aim of this study is to carry out a preliminary study for the adaptation and validation of a computerized task for the recognition of complex emotions and mental states in dynamic facial expressions of the CAM-C test, analyzing its psychometric properties in Argentine children and adolescents from 9 to 14 years old by measuring hit rates and reaction times.

Method

Participants

The pilot study comprised 35 children and adolescents (61.5% girls and 28.2% boys) from Buenos Aires (Argentina), aged 9 to 14 (M = 11.29, SD = 1.56). Then, the normative sample comprised 135 children and adolescents (67.4% girls and 32.6% boys) from Buenos Aires (Argentina), also aged 9 to 14 (M = 11.57, SD = 1.41). Of the total sample, 77.5% of the participants studied in public institutions and 22.5% in private institutions. The racial/ethnic composition of the sample was Hispanic/Latin White. The participants were selected by non-probabilistic and intentional sampling, so the results obtained from the sample cannot be generalized to the total population, because the obtained sample does not represent the community. The participants carried out individually the task in a classroom of the educational institution, supervised by qualified professionals.

During the development of the study, the ethical principles of research with human beings were followed, ensuring the necessary conditions to protect the anonymity and confidentiality of the data. Participation required informed consent from parents and assent from children. Children and adolescents with a confirmed diagnosis of neurodevelopmental disorders were excluded, according to DSM-5 (American Psychiatric Association, 2013), since the objective of this study is to provide evidence to adapt and validate the test to a normotypical population of Argentine children.

The research was approved by the Bioethics Committee of the National University of Mar del Plata, registered in the Provincial Registry of Research Ethics Committees reporting to the Central Research Ethics Committee – Ministry of Health of the Province of Buenos Aires.

Instruments

Sociodemographic Questionnaire

It is a self-administered questionnaire built ad hoc to obtain sociodemographic data of the sample, such as sex, age, and place of residence.

The Cambridge Mindreading Face-Voice Battery for Children (CAM-C)

The CAM-C is based on the adult version of the same instrument (Golan et al., 2006, 2015). This battery tests recognition of nine complex emotions and mental states (loving, embarrassed, undecided, unfriendly, bothered, nervous, disappointed, amused, and jealous) in children older than 8 years (M = 10.0, SD = 1.1), using two unimodal tasks: a face task, comprising silent video clips of child and adult actors, expressing the emotions on their faces, and a voice task, comprising recordings of short sentences expressing various emotional intonations. The selected concepts included emotions that are developmentally significant, subtle variations of basic emotions that have a mental component, and emotions and mental states that are important for the everyday social functioning. All stimuli were taken from Mind Reading (Baron-Cohen et al., 2004). In this study, only the face recognition test is taken for its adaptation and validation.

The psychometric properties of the test show good reliability, showing acceptable correlations in test-retest (r = .74, p < .001), and evidence of concurrent validity: the CAM-C face task was negatively correlated at a significant level with the Childhood Autism Spectrum Test (CAST; r = -.54, p < .001) and positively with the Reading the Mind in the Eyes Child Version (RME; r = .35, p < .001). Likewise, age was also positively correlated with the CAM-C face task (r = .53, p < .001) (Golan et al., 2015). The CAM-C effectively discriminates between children with high-functioning autism spectrum disorder (HFASD) and typical children and has been recommended as a standardized measure to be used in ER studies for children with HFASD (Thomeer et al., 2015).

New Spanish Empathy Questionnaire for Children and Early Adolescents

This scale consists of 15 items that assess empathy, in a multidimensional way, based on the social cognitive neuroscience model. It is designed to evaluate empathy in Argentine children through five dimensions: emotional contagion (items 1, 5, 8), self-awareness (items 10, 12, 15), perspective taking (items 3, 6, 14), emotional regulation (items 4, 7, 11, all inverse) and empathic attitude (items 2, 9, 13). This scale was developed and validated in Argentina and has adequate psychometric properties (Richaud et al., 2017). The internal consistency for each factor is acceptable (ω = .75 for emotional contagion, ω = .76 for self-awareness, ω = .72 for perspective taking, ω = .72 for emotional regulation, and ω = .70 for empathic attitude) and the five factors correlate appropriately with criteria variables, such as prosocial behavior, emotional regulation, emotional instability, aggression, and perspective taking (IRI) (Richaud et al., 2017).

Procedure

The procedure followed has contemplated the international professional regulations for the adaptation and validation of tests used in clinical and institutional practice (American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014; American Psychological Association, 2010) and, more specifically, in psychological research (International Test Commission, 2014).

Authorization was first requested and obtained from the main author of the test. Due to the original language of the test is English, the translation of the task was carried out in order to create a version that corresponded to the local language and the linguistic styles of the context in which it was applied. A linguistic and conceptual evaluation of the meaning of the items was carried out considering the terminology that best suited our cultural context and, at the same time, the aspects assessed by the test. Two forward- and two back-translations were done in parallel by translating psychologists. This method was used to ensure the translated version would be grammatically sound and the terms used were correct. After the reconcilement of the two forward and back translations, sentence revision was carried out. In order to culturally adapt the translated version for the Argentine child population, a panel of ten experts in measuring emotional variables and psychometrics reviewed the first translated version. They were asked to judge each item considering its formal quality (semantic clarity, syntactic correctness, and suitability for the target population) and to make all the necessary observations and suggestions in order to improve the task. The degree of agreement between judges was evaluated by calculating the Aiken V index, considering as a criterion that at least 70% of the judges agreed that the content of the item was relevant and effectively represented the video to which it belonged. The result was the final version of the task, ready for field testing.

A software was implemented with the purpose of accurately registering the responses and reaction times, which allowed to unify and systematize the variables of presentation of the stimuli, minimizing the possible differences that could arise between evaluators and guaranteeing the reliability of the measurement.

Then, a pilot study was carried out, allowed the linguistic adaptation, and provided evidence of face validity, since the participants were asked to contribute their opinions to improving the format and content of the items and investigating the effectiveness of the emotional stimuli.

Based on the results, the corresponding modifications were made and the final version of the instrument was created. The inclusion criteria for each video indicated in the original test were met: items were included if the target answer was picked by at least half of the participants and if no foil was selected by more than a third of the participants (p < .05, binomial test). This final version of the task was administered to the normative sample, together with the empathy questionnaire.

Figure 1. An Item Example from the Face Task (showing one frame of the full video clip).
Note. Image retrieved from *Mindreading: The interactive guide to emotion.* Courtesy of Jessica Kingsley Ltd.

Participants were individually tested at a local school. Tasks were presented to the participants on a laptop computer with a 15-in. screen. Following the original test procedure, for each emotional concept, three silent video clips of facial expressions were used (each video lasted approximately 5 seconds); 29 faces were portrayed by

professional actors, both male and female, of different age groups and ethnicities (15 videos represented by men and 14 by women, 9 videos represented by children and 20 by adults, 23 videos represented by white people, and 6 by nonwhite people). This task was carried out using experimental software, starting with an instruction slide, asking participants to choose the answer that best described how the person in each clip was feeling. The instructions were followed by two practice items. The videos were presented sequentially and randomly, interspersing the nine emotions. Due to the difficulty observed in the pilot study in relation to the children's reading speed, it was decided that the four emotion labels, numbered from 1 to 4, were presented before playing each clip. So, once the participants had read the four response options, the video was played. The four response options consisted of the target emotion/mental state (i.e., the emotion/mental state that the actor intended to express), and three control emotions/mental states. An example from the Face Task is offered in Figure 1.

In addition, to avoid confusing effects due to reading difficulties, the administrator read the instructions and the response options to the children, and used the computer commands, selecting the response chosen verbally by the child. After choosing an answer, the next item was presented. No feedback was given during the task.

Regarding the score, one point was assigned to each hit and zero points to each error, constituting a scale of minimum value 0 and maximum value 27. Likewise, the measurement of the reaction times used for the recognition of each emotion was added to the original test. Reaction times were measured by the software through the exact recording of the time elapsed between the application of a video and the beginning of participants' response. There was no time limit to answer each item. Completion of the whole battery took about 20 min, including breaks. The Empathy Questionnaire was completed during the same session, taking about 10 min.

Data Analysis

In the first place, a content validity study was carried out with the aim of assessing both the formal aspects of the test (assignment, difficulty, theoretical dimension, quantity, and quality of the items) and the videos. To do that, the Aiken coefficient V was estimated, which can vary between 0 and 1, and must reach at least a critical value V = .50 to be considered acceptable according to the criteria established by Aiken (1985). However, more recent studies suggest that more conservative levels be considered (V values ≥ .70) and pay attention to the confidence intervals of the coefficient (Soto & Segovia, 2009). To estimate the coefficient and its confidence intervals, the program developed by Soto and Segovia (2009) was used. Taking into consideration the recommendations by Soto and Segovia, it was established as a criterion that the lower limit of the intervals obtained should be values equal to or greater than .70.

Subsequently, the ViSta version 7.2.04 software was used to estimate the difficulty (p) and discrimination indices. For the interpretation of *p*, the values were considered very easy (.81 to .100), easy (.61 to .80), moderately easy (.41 to .60), difficult (.21 to .40), and very difficult (.01 to .20). To evaluate the discriminative power of the videos (their ability to distinguish between those who show high and low levels in the criterion) the discrimination index (d) was used. For d, values >.39 were considered excellent, .30 to .38 good, .20 to .29 regular, .00 to .20 poor, and < -.01 worst (Ebel & Frisbie, 1986). Considering that d tends to be biased towards intermediate degrees of difficulty, the point biserial correlation (R_{bp}) was used, which is a measure of the relationship between the item and the criterion, independently from the difficulty of the item (Guilford & Frutchter, 1978). Values >.35 are considered with excellent discriminative power, .26 to .35 good, .15 to .25 regular, .00 to .14 poor, and < .00 with negative discrimination (Diaz Rojas & Leyva Sánchez, 2013).

To evaluate the internal structure of the CAM-C FACE, an exploratory factor analysis (EFA) was performed using Factor statistical software version 11.05.01, using unweighted least squares (ULS) and oblimin rotation (direct oblimin) as the estimation method. Since no studies have been reported yet that analyze the factorial structure of the CAM-C FACE test, in this study two factor solutions were evaluated: a two-factor solution and a one-dimensional solution. Kaiser-Meyer-Olkin (KMO) sample adequacy indices and the Bartlett's sphericity test were used to assess the feasibility of factor analysis. Additionally, the goodness of fit index (GFI), and the root mean square of the residuals (RMSR) were considered. GFI values ≥ .95 and close to 0 for RMSR are indicators of a good model fit. A critical value of .30 was considered for factor saturation (Lloret et al., 2014). For conducting the EFA, the hit rates (right or wrong responses) of the participants and the reaction times to each of the videos were considered.

Moreover, the reliability of the test was estimated, using the Kuder-Richardson 20 coefficient (KR20) for correct answers, and the composite reliability (ρ) for reaction times. Values equal to or greater than .70 for both coefficients were considered acceptable (Nunnally, 1978).

Regarding sociodemographic variables, the existence of differences was analyzed according to sex and age by the Student's test. Additionally, a series of bivariate correlations using Pearson's coefficient was done in order to examine the relationship of the CAM-C FACE with the New Empathy Questionnaire in Spanish for Children and Adolescents (convergent validity). Emotion recognition test scores are often reported to correlate with self-reported empathy (Vellante et al., 2013). Finally, the effect size was estimated using the Cohen's *d* statistic. For their interpretation, values *d* = .20 were considered small, *d* = .50 medium, and *d* = .80 large (Cohen, 1988). This study was not pre-registered. Data and study materials are not available to other researchers.

Results

Linguistic Adaptation and Pilot Study: Content Validity and Face Validity

Study by Expert Judges

A backward translation was performed with the aim of generating a certain level of equivalence between the original and the translated version. In addition, modifications were made to adapt linguistic styles of the local context in which it was applied. The results of the study of judges (*N* = 12) indicated that, although most of the V coefficients exceed the critical value of .70, corresponding to the lower limit of the confidence interval, items (videos) 14, 15, 19, 20, and 22 did not meet this criterion (see Table 1). Regarding the formal aspects of the test, all met the critical value of the lower limit of the confidence interval: instruction difficulty (V coefficient = .813 [.704, .888]), appropriate number of videos (V coefficient = .958 [.882, .986]), and appropriate video quality (V coefficient = .813 [.704, .888]).

Pilot Study

From the results obtained, item translations were modified taking into account the observations made by the judges. For this, a series of cognitive interviews were conducted with 35 children and adolescents from two schools Buenos Aires (Argentina) in order to adapt the vocabulary of the statements to the target population, collect information on possible content or format errors, and ensure that items and instructions were understood correctly (Caicedo & Zalazar, 2018).

Table 1. Agreement between Judges of the CAM-C FACE Test Videos

Emotions	Videos	Aiken's V [90% IC]
Antipático/a (unfriendly)	Video 1	.917 [.827, .962]
	Video 2	.833 [.728, .903]
	Video 3	.813 [.704, .888]
Indeciso/a (undecided)	Video 4	.896 [.801, .948]
	Video 5	.938 [.854, .975]
	Video 6	.958 [.882, .986]
Nervioso/a (nervous)	Video 7	.833 [.728, .903]
	Video 8	.938 [.854, .975]
	Video 9	.854 [.752, .919]
Amoroso/a (loving)	Video 10	.833 [.728, .903]
	Video 11	.854 [.752, .919]
	Video 12	.896 [.801, .948]
Celoso/a (jealous)	Video 13	.833 [.728, .903]
	Video 14	.563 [.445, 674][1]
	Video 15	.688 [.570, .785][1]
Avergonzado/a (embarrassed)	Video 16	.854 [.752, .919]
	Video 17	.833 [.728, .903]
	Video 18	.854 [.752, .919]
Decepcionado/a (disappointed)	Video 19	.750 [.636, .838] [1]
	Video 20	.729 [.614, .820] [1]
	Video 21	.854 [.752, .919]
Preocupado/a (bothered)	Video 22	.771 [.658, .855] [1]
	Video 23	.875 [.776, .934]
	Video 24	.896 [.801, .948]
Divertido/a (amused)	Video 25	.917 [.827, .962]
	Video 26	.813 [.704, .888]
	Video 27	.854 [.752, .919]

Note. [1]Does not meet criteria.

Recognition indices for each video were evaluated, calculating the percentage of correct answers and the percentage for each foil. Following the original test procedure, the targets that met the following criteria were kept: the correct answer should be picked by at least half of the participants, while the foil should not be selected by more than a third of the participants ($p < .05$, binomial test) (Table 2).

As can be seen, videos 3, 9, 12, and 24 did not meet the inclusion criteria indicated in the original test, so foil greater than .33 were replaced by others. In video 3, the foil used was *disgustada (*.46) and was replaced by *aburrida*; in video 9, the foil used was *irritada* (.34) and was replaced by *enojada*; in video 12, the foil used was *esperanzada* (.34) and was replaced by *educada*; and in video 24, the foil used was *desconfiado* (.34) and was replaced by *furioso*. The change was made by searching for an alternative term among those already present in the different exercises of the test, which would not generate confusion with the correct option, but which would retain a similar emotional valence (positive vs. negative).

Regarding the contributions and opinions requested from the participants to improve the format and content of the items and investigate the effectiveness of the emotional stimuli, it was observed that the participants did not understand the term *antipático/a* [unfriendly]. Therefore, it was replaced by *mala onda,* a term taken from participants' suggestions.

Psychometric Properties of the Argentine Version of the CAM-C FACE Test

To study different aspects of facial emotion recognition, two valid methods were analyzed: record of hit rates (the number of right answers) and reaction times of all answers. Therefore, accuracy and performance speed can be evaluated, respectively, providing a better differentiation between participants and between stimuli (Kosonogov & Titova, 2019).

Table 2. Pilot Test Recognition Rates of the CAM-C FACE Test Videos

Emotions	Videos	Recognition Indices			
		Correct Answer	Foil 1	Foil 2	Foil 3
Antipático/a (unfriendly)	Video 1	.69	.17	.11	.03
	Video 2	.69	.20	.09	.03
	Video 3	.29[1]	.46[1]	.23	.03
Indeciso/a (undecided)	Video 4	.83	.09	.06	.03
	Video 5	.89	.11	.00	.00
	Video 6	.77	.20	.03	.00
Nervioso/a (nervous)	Video 7	.74	.23	.03	.00
	Video 8	.86	.09	.03	.03
	Video 9	.63	.34[1]	.03	.00
Amoroso/a (loving)	Video 10	.77	.11	.09	.03
	Video 11	.94	.03	.03	.00
	Video 12	.63	.34[1]	.03	.00
Celoso/a (jealous)	Video 13	.66	.20	.11	.03
	Video 14	.66	.29	.03	.03
	Video 15	.54	.29	.11	.06
A*vergonzado/a* (embarrassed)	Video 16	.69	.14	.11	.06
	Video 17	.54	.23	.17	.06
	Video 18	.71	.14	.11	.03
Decepcionado/a (disappointed)	Video 19	.63	.26	.11	.00
	Video 20	.57	.26	.14	.03
	Video 21	.80	.17	.03	.00
Preocupado/a (bothered)	Video 22	.63	.23	.09	.06
	Video 23	.71	.20	.09	.00
	Video 24	.26[1]	.60[1]	.11	.03
Divertido/a (amused)	Video 25	.80	.11	.06	.03
	Video 26	.74	.11	.11	.03
	Video 27	.89	.06	.03	.03
Totals		.71	.19	.08	.02

Note. [1]Does not meet criteria.

Table 3. Proportions, Variances, Difficulty Index (p), Discrimination (d), and Biserial Correlation (R_{bis}) of Correct Answer of the CAM-C FACE

Emotion	Videos	Proportions	Variances	p	d	R_{bis}
	Video 1	85.2	.127	.851	.15	.385
Mala onda (unfriendly)	Video 2	77.0	.178	.770	.23	.271
	Video 3	80.0	.161	.800	.20	.256
	Video 4	65.2	.229	.651	.35	.443
I*ndeciso/a* (undecided)	Video 5	77.0	.178	.770	.23	.564
	Video 6	65.2	.229	.651	.35	.439
	Video 7	83.0	.142	.829	.17	.474
Nervioso/a (nervous)	Video 8	82.2	.147	.822	.18	.442
	Video 9	60.0	.242	.600	.40	.307
	Video 10	54.1	.250	.540	.46	.165
Amoroso/a (loving)	Video 11	69.6	.213	.696	.30	.186
	Video 12	54.1	.250	.540	.46	.289
	Video 13	61.5	.239	.614	.39	.140
Celoso/a (jealous)	Video 14	48.9	.252	.488	.51	.353
	Video 15	58.5	.245	.585	.41	.366
	Video 16	74.1	.193	.740	.26	.329
Avergonzado/a (embarrassed)	Video 17	63.7	.233	.637	.36	.389
	Video 18	61.5	.239	.614	.39	.366
	Video 19	60.0	.242	.600	.40	.235
Decepcionado/a (disappointed)	Video 20	53.3	.251	.533	.47	.334
	Video 21	54.8	.250	.548	.45	.225
	Video 22	65.2	.229	.651	.35	.318
Preocupado/a (bothered)	Video 23	69.6	.213	.696	.30	.352
	Video 24	61.5	.239	.614	.39	.289
	Video 25	77.8	.174	.777	.22	.139
Divertido/a (amused)	Video 26	71.9	.204	.718	.28	.276
	Video 27	71.9	.204	.718	.28	.351

Item Analysis of Hit Rates. Assessing the Accuracy of Performance

No missing cases were reported because participants were tested individually by the researcher, who notified them if any item was left unanswered.

For this analysis, proportions, variances, difficulty index, discrimination, and biserial correlation of correct answer were measured, as it is a dichotomous variable (see Table 3). Regarding item difficulty (p), it is observed that all videos (items) were moderately easy to very easy (≥ .41), with a higher proportion of easy videos. No video turned out to be difficult or very difficult, reaching a low number of correct answers (see Table 4).

Table 4. Difficulty Index (p) Categorized for the Correct Answer of the CAM-C FACE

Item Classification	Difficulty Index Value	Videos	%
Very easy	.81-1.00	1, 7, 8	11.1
Easy	.61-.80	2, 3, 4, 5, 6, 11, 13, 16, 17, 18, 22, 23, 24, 25, 26, 27	59.3
Moderately easy	.41-.60	9, 10, 12, 14, 15, 19, 20, 21	29.6
Hard	.21-.40	-	-
Very difficult	.01-.20	-	-

Regarding the discrimination index, values between .15 (video 1) and .51 (video 14) were observed. According to Ebel and Frisbie's (1986) classification, 11 items (40.7%) presented an excellent discriminant value, 6 items presented a good value (22.2%), 7 items had a fair value (25.9%), and 3 items showed a poor value (11.1%) (see Table 5). Although videos 2, 3, 5, 16, 25, 26, and 27 deserve attention due to their low discriminative power, more attention deserves videos 1, 7, and 8, which should be dismissed or reviewed in depth.

Table 5. Classification of the Videos of the Correct Answer of the CAM-C FACE Test according to their *d* Value

Item Classification	Valor *d*	Videos	%
Excellent	> .39	9, 10, 12, 13, 14, 15, 18, 19, 20, 21, 24	40.7
Good	.30-.38	4, 6, 11, 17, 22, 23	22.2
Regular	.20-.29	2, 3, 5, 16, 25, 26, 27	25.9
Poor	.00-.20	1, 7, 8	11.1
Lousy	< .01	-	-

Regarding the distribution of the results of the biserial point correlation (r_{bp}), the largest number of videos presented excellent (37%) and good (40.7%) discriminative power. Then, some videos presented a regular (14.8%) and poor (7.4%) discriminative power (see Table 6). Based on the results presented and considering the preliminary characteristics of the study, it was decided that videos should be retained despite their ease and low discriminating power.

Item Analysis of Reaction Times. Assessing the Speed of Performance

Furthermore, a descriptive analysis of reaction times was performed. In this analysis, mean, standard deviation skewness, and kurtosis of reaction times were measured, as it is a continuous variable.

Regarding asymmetry, it was observed that 8 items presented values between +1 and -1, 14 items presented values between +2 and -2, and 5 items presented values > ±2.00. Regarding kurtosis, 8 items presented values between +1 and -1, 2 items presented values between ±2.00, and 17 items presented values > ±2.00 (see Table 7).

Table 6. Classification of the Videos of the Correct Answer of the CAM-C FACE Test according to the Biserial Correlation Point Discrimination Coefficient

Item Classification	Valor r_{bp}	Videos	%
Excellent	> .36	1, 4, 5, 6, 7, 8, 15, 17, 18, 27	37
Good	.26-.35	2, 3, 9, 12, 14, 16, 20, 22, 23, 24, 26	40.7
Regular	.15-.25	10, 11, 19, 21	14.8
Poor	.00-.14	13, 25	7.4
Negatively	< .00	-	-

Table 7. Mean, Standard Deviation, Skewness, and Kurtosis of Reaction Times of the CAM-C FACE

Emotion	Videos	*M*	*SD*	Asymmetry	Kurtosis
Mala onda (unfriendly)	Video 1	5.008	1.702	0.494	-0.164
	Video 2	5.134	1.574	0.282	0.084
	Video 3	4.824	1.808	0.967	1.488
Indeciso/a (undecided)	Video 4	5.537	2.273	1.798	6.713
	Video 5	4.491	2.232	1.446	3.887
	Video 6	4.726	2.330	2.448	11.060
Nervioso/a (nervous)	Video 7	4.622	1.617	0.788	0.799
	Video 8	5.191	2.305	4.557	38.152
	Video 9	5.233	2.024	1.447	4.871
Amoroso/a (loving)	Video 10	5.254	2.530	3.009	17.841
	Video 11	5.145	2.416	1.085	0.617
	Video 12	4.647	2.117	1.441	4.269
Celoso/a (jealous)	Video 13	5.046	2.172	1.139	2.156
	Video 14	6.018	2.345	0.909	1.050
	Video 15	5.014	1.719	0.467	0.168
Avergonzado/a (embarrassed)	Video 16	5.816	2.424	1.680	6.562
	Video 17	5.057	1.875	2.182	10.770
	Video 18	5.188	1.952	0.484	-0.694
Decepcionado/a (disappointed)	Video 19	4.925	1.738	0.498	0.004
	Video 20	5.339	2.224	1.555	4.462
	Video 21	4.997	2.069	1.332	2.765
Preocupado/a (bothered)	Video 22	6.124	2.881	1.849	4.999
	Video 23	5.124	1.883	1.047	2.415
	Video 24	5.720	2.464	1.796	5.377
Divertido/a (amused)	Video 25	4.685	2.009	1.394	3.385
	Video 26	4.013	1.925	2.935	13.429
	Video 27	4.976	2.406	1.031	0.459

Evidence of Internal Structure

When taking into account the number of correct answers of the participants, the measure of sample adequacy, the measure of the two-factor model, the KMO obtained (.539), and Bartlett's sphericity test with values of 538.5 (df = 351, p = .000) for the two-factor and one-factor models indicated that it is not possible to apply factor analysis. An inspection of these results could lie in the restriction to the range in the answers of the participants, that is, the low variability of the answers due to the high prevalence of items of low difficulty.

Moreover, when exploring the factorial structure according to reaction times, the KMO sample adequacy measure obtained (.928) and Bartlett's sphericity test with values of 1428.5 (df = 351, p = .000) suggest that it is possible to apply factor analysis for both models. Regarding the two-factor model, it explained 48% of the variance. The fit indices obtained were satisfactory (GFI = .987, RMSR = .052). However, inspection of factor loadings suggests that a one-dimensional structure is more appropriate due to the distribution of factor loadings. Indeed, only two videos (4 and 26) were observed to present contributions in factor 2 (see Table 8). The one-dimensional model explained 44% of the variance, whose fit indices (GFI = .985, RMSR = .057) were satisfactory. The inspection of factor loads ranged between .312 (video 26) and .718 (video 19).

Table 8. Factor Loads, Two-factor and One-dimensional Model, according to Reaction Times from CAM-C FACE

		Solution		
Emoción	Videos	Factor 1	Factor 2	Unidimensional
Mala onda (unfriendly)	Video 1	**.695**	-.069	**.654**
	Video 2	**.657**	.078	**.697**
	Video 3	**.655**	-.018	**.643**
Indeciso/a (undecided)	Video 4	.436	**.532**	**.710**
	Video 5	**.605**	.139	**.678**
	Video 6	**.606**	.040	**.625**
Nervioso/a (nervous)	Video 7	**.504**	.338	**.680**
	Video 8	**.531**	.141	**.604**
	Video 9	**.661**	-.010	**.654**
Amoroso/a (loving)	Video 10	**.713**	-.232	**.582**
	Video 11	**.698**	-.055	**.665**
	Video 12	**.749**	-.052	**.717**
Celoso/a (jealous)	Video 13	**.606**	.145	**.681**
	Video 14	**.652**	.102	**.704**
	Video 15	**.676**	.007	**.677**
Avergonzado/a (embarrassed)	Video 16	**.555**	.181	**.650**
	Video 17	**.487**	.078	**.527**
	Video 18	**.638**	.073	**.675**
Decepcionado/a (disappointed)	Video 19	**.721**	.000	**.718**
	Video 20	**.689**	-.024	**.673**
	Video 21	**.766**	-.193	**.656**
Preocupado/a (bothered)	Video 22	**.604**	.009	**.607**
	Video 23	**.540**	.099	**.591**
	Video 24	**.558**	.070	**.594**
Divertido/a (amused)	Video 25	**.653**	-.077	**.608**
	Video 26	.045	**.506**	**.312**
	Video 27	**.578**	.216	**.691**

Internal Consistency

Although EFA based on correct answers was not feasible, reliability was moderate (KD-20 = .660). On the other hand, when considering the reaction times, the values obtained were satisfactory for the one-dimensional CAM-C FACE model (ρ = .950).

Relationship with Sociodemographic Features

Once the assumptions of normality have been confirmed and homoscedasticity through the tests of Kolmogorov-Smirnov and Levene proceeded to perform the *t*-test.

The results indicated that girls presented more correct answers than boys, M_{wom} = 18.620, SD_{wom} = 3.907; M_{men} = 17.000, SD_{men} = 3.916; $t(133)$ = -2.279, p = .024, with a medium effect size (d = 0.414), while boys had longer reaction times, M_{wom} = 133.452, SD_{wom} = 38.353; M_{men} = 146.405, SD_{men} = 34.418; $t(133)$ = 1.924, p = .056, with a small effect size (d = 0.355). Regarding age, the sample was divided into two groups (group 1= from 9 to 11 years old, group 2= from 12 to 14 years old) and the results indicated that group 2 presented more correct answers compared to group 1, M_1= 16.790, SD_1 = 4.164; M_2 = 19.290, SD_2 = 3.374; $t(133)$ = -3.843, p = .000, with a medium effect size (d = 0.659), while no differences were observed between groups in terms of reaction times, M_1 = 137.430, SD_1 = 40.534; M_2 = 138.283, SD_2 = 34.517; $t(133)$ = -.132, p = .895.

Evidence of Criterion Validity

Regarding convergent validity, a direct correlation of large effect was observed between correct answers of the CAM-C FACE and total

score of the Empathy Questionnaire (r = .418, p < .000, d = 0.923), and a negative correlation of large effect between reaction times of the CAM-C FACE and total score of the Empathy Questionnaire (r = -.224, p < .009, small effect, d = 0.871).

Discussion

This preliminary study reports the results of a software development of the Argentine version of the Cambridge Mindreading Face-Voice Battery for Children (CAM-C). A software was implemented with the purpose of accurately registering the responses and reaction times, so two dependent variables were analysed: record of hit rates (the number of right answers) and reaction times of all answers. This method allows assessing the accuracy and the speed of performance, respectively, providing a better differentiation between participants and between stimuli (Kosonogov & Titova, 2019).

When assessing the accuracy of performance through item analysis of hit rates, results show that all videos were classified as very easy to moderately easy (reaching a large number of correct answers) and none was difficult for this Argentine age population. This low variability of the answers, due to the high prevalence of items of low difficulty, shows the range restriction in the participants' answers, which is a limitation of the test for this age range.

Therefore, it is observed that, since the task has not a wide score range, it presents a ceiling effect that fail to distinguish different levels of performance, even taking into account that it is considered that the use of video clips tends to lead to smaller standard deviations and greater differences (Kosonogov & Titova, 2019). However, this is considered to be expected when evaluating basically cognitive processing because the probability of making an error is low, but also the ceiling effect can be a consequence of the characteristic of the population, like a specific age cohort (Schweizer et al., 2019). Although the test was designed in order to evaluate the emotional recognition of children with autism, it is observed that it will be necessary to take precautions for its use in a typical population, if only the number of correct answers is taken into account. In addition, the restriction of range and its consequences may explain some of the inconsistencies and contradictory findings in the field, like age or gender differences (Vaci et al., 2014).

On the other hand, it has been proposed that the degree of difficulty of an item might vary from the original to the translated version because of cultural differences in the meaning attributed to the target definitions, but this hypothesis can be tested by a comparative study only (Hallerbäck et al., 2009).

Beside this, according to Backhoff et al. (2000), a test should have 5% easy items, 20% moderately easy, 50% medium difficulty, 20% moderately difficult, and 5% difficult, which is not fulfilled in this test for this age population.

Moreover, results show that almost 80% of videos presented good discriminative power. This means that most videos have a good ability to distinguish between people with high and low values in the test. Nevertheless, videos 10, 11, 19, 21, 13, and 25 should be discarded or reviewed in depth. Since they are not recognized even by those participants who present a greater capacity to recognize emotions, it can be assumed that those videos show a poor quality of the emotional expression.

In relation to the evidence of internal structure, due to the range restriction in the answers, it was not possible to apply factor analysis considering the number of correct answers (accuracy of performance). However, when the factorial structure is explored according to reaction times (speed of performance), inspection of factor loadings suggests that a one-dimensional structure is more appropriate due to the distribution of factor loadings, explaining 44% of the variance, whose fit indices (GFI = .985, RMSR = .057) were satisfactory. The inspection of the factor loads ranged between .312 (video 26) and .718 (video 19). The unidimensional model, assuming all items loaded on a single factor, was observed in previous studies that analyze the factorial structure of a similar test, widely studied and used in the field of emotional recognition: the Eyes Test (Vellante et al., 2013). However, no studies have been reported that analyze the factorial structure of the CAM-C FACE test.

Despite being a small sample and even being an easy test for the target age, it is observed that reaction times would allow establishing individual differences based on the speed of performance, so this indicator is perhaps more important to take into account when Argentine children of these ages are evaluated.

Although EFA based on correct answers was not feasible, reliability was moderate (KD-20 = .660), which would indicate that some items are not measuring the construct of interest. In previous studies, test-retest reliability was examined (Golan et al., 2015; Rodgers et al., 2021) and internal consistency was calculated using Cronbach's alpha (Rodgers et al., 2021; with α = .72). However, as the literature indicates, this coefficient is not adequate for dichotomous items (Martinez-Arias et al., 2014). On the other hand, when considering reaction times, the values obtained were satisfactory for the one-dimensional CAM-C FACE model (ρ = .950).

Regarding sociodemographic features, results indicated that girls presented more correct answers than boys, while boys had longer reaction times. Given that, methodologically, reaction time is considered to be one of the experimental procedures that provides the most reliable quantitative data for the study of mental processes, females have been observed to show more accurate facial emotion recognition compared to males and were faster in correctly recognizing facial emotions (Wingenbach et al., 2018).

In terms of age, results indicated that the group of children from 12 to 14 years old presented more correct answers compared to the 9-to-11-year-old group, while no differences were observed between groups in terms of reaction times. This agrees with what was reported by Golan et al. (2008) about the correlation between task scores and age, indicating that the ability to recognize complex emotions and mental states improves with age. These results are also in agreement with the study carried out by Rodgers et al. (2021), whose results revealed significant positive correlations between CAM-C Faces scores and child age, but there were no significant associations between CAM-C Faces scores and sex.

Finally, the results show, as expected, a direct correlation of large effect between correct answers of the CAM-C FACE and total score of the Empathy Questionnaire, and a negative correlation of large effect between reaction times of the CAM-C FACE and total score of the Empathy Questionnaire, which provides evidence of convergent validity. Typically, correlations are obtained between emotional recognition and empathy (Vellante et al., 2013).

Even though, to our knowledge, this is the first study to assess all psychometric properties of the CAM-C FACE test, additional studies are needed to confirm the results.

Implications

Having locally developed tests for the facial recognition of complex emotions in the area of neuropsychology has great clinical implications, since the alteration of this ability has been reported in multiple pathologies. Likewise, this skill constitutes a basic process in the emotional and social skills that are so important in school-age children.

Limitations and Future Directions

The results from the present study should be interpreted in the light of limitations that suggest potential areas for future research.

First, the size of the sample is small. Guidelines recommend a minimum EFA sample size of at least 300 participants (Field, 2013;

Tabachnick & Fidell, 2013), so future work using larger samples will be an important next step in continued testing of the psychometric properties of this measure.

Second, in the sample only children and adolescents aged 9 to 14 years old are included. The lack of more varied strata of the population precludes the possibility of comparing the performance of this population with younger children. Another limitation involves the lack of a comparison group to determine whether the CAM-C FACE accurately differentiates children with and without a clinical disorder. This limits generalizability of the findings and indicates a need for more testing with diverse samples and with a wider range of cognitive and language abilities in order to establish the sensitivity and specificity of the instrument to detect deficits, and perhaps even talents. Further, the current study only compared CAM-C FACE scores and reaction times to empathy self-report and future studies using observations would provide important information on how emotion recognition skills might be linked to actual social behaviors.

Another limitation of the study consists of the way of recording reaction times in emotional recognition. As the latency of the response was recorded by a motor action of the experimenter on the computer screen, the recordings could be slightly inaccurate, being altered by the sensory abilities of the experimenter in the detection of the voice and the subsequent execution of a motor response to stop recording reaction times. Subtle millisecond differences could provide less accurate data that could limit the scope of the conclusions. For future studies, these effects could be attenuated by software for recording the vocal response (voice key).

Finally, one limitation of the test is that it does not offer the possibility of studying the differences in the accuracy of recognition between emotions because the number of trials is very small (three of each emotion).

Beyond the limitations observed, as mentioned, this is the first study providing evidence that this test can be scored as a single factor when assessing the speed of performance through reaction times. Having an instrument to evaluate the recognition of complex emotions in Argentine children and adolescents is imperative because, although there are tests worldwide to evaluate emotional recognition of non-basic emotions at these ages, these tests have not been adapted and validated in our country and, as they come from different languages and cultures, they lose validity and consistency when applied to the local context, mainly because emotional phenomena are modulated by situational and cultural factors (Medrano et al., 2013). Therefore, this study may represent the beginning of the provision of an instrumental response to a vacant area in our context. After further research, this test may be useful to assess differences in recognition of complex affective states and in intervention research to monitor improvements in this skill, or to improve diagnostic assessments.

Conflict of Interest

The authors of this article declare no conflict of interest.

References

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology, 12*(2), 169-177. https://doi.org/10.1016/S0959-4388(02)00301-X

Adolphs, R., Tranel, D., & Damasio, A. R. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition, 52*(1), 61-69. https://doi.org/10.1016/S0278-2626(03)00009-5

Aiken, L. R. (1985). Three coeficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*, 131-142.

American Educational Research Association, American Psychological Association and National Council on Measurement in Education (2014). *Standards for educational and psychological testing.* AERA.

American Psychological Association. (2010). *Ethical principles for psychologists and code of conduct.* American Psychiatric Association.

American Psychological Association. (2013). Guidelines for psychological evaluations in child protection matters. *The American Psychologist, 68*(1), 20-31. https://doi.org/10.1037/a0029891

Backhoff, E., Larrazolo, N., & Rosas, M. (2000). The level of difficulty and discrimination power of the Basic Knowledge and Skills Examination (EXHCOBA). *Revista Electrónica de Investigación Educativa, 2*(1), 1-16. http://redie.uabc.mx/vol2no1/contents-backhoff.html

Barceló-Martínez, E., Fonseca-Consuegra, L., Aguirre-Acevedo, D. C., Gelves-Ospina, M., Roman, N. F., & Benítez-Agudelo, J. C. (2018). Datos normativos del Cambridge Mindreading test en español en adultos jóvenes de la ciudad de Barranquilla, Colombia. *Archivos de Neurociencias, 23*(1), 7-15. https://doi.org/10.31157/archneurosciencesmex.v23i1.1

Baron-Cohen, S., Golan, O., Wheelwright, S., & Hill, J. J. (2004). *Mind reading: The interactive guide to emotions.* Jessica Kingsley Limited. www.jkp.com

Brothers, L., & Ring, B. (1992). A neuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience, 4*(2), 107-118. https://doi.org/10.1162/jocn.1992.4.2.107

Caicedo Cavagnis, E., & Zalazar-Jaime, M. F. (2018). Entrevistas cognitivas: revisión, directrices de uso y aplicación en investigaciones psicológicas. *Avalição Psicológica, 17*(3), 45-67. https://doi.org/10.15689/ap.2018.1703.14883.09

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Damasio, A. R. (2010). *El error de Descartes: la razón de las emociones.* Editorial Paidos.

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: Speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology, 24*(2), 200-213. https://doi.org/10.1076/jcen.24.2.200.989

Díaz Rojas, P. A., & Leyva Sánchez, E. (2013). Metodología para determinar la calidad de los instrumentos de evaluación. *Educación Médica Superior, 27*(2), 269-286.

Ebel, R. L., & Frisbie, D. A. (1986). *Essentials of educational measurement* (4th ed.). Prentice-Hall, Inc.

Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clinical Psychology Review, 22*(6), 789-832. https://doi.org/10.1016/S0272-7358(02)00130-7

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic? *Emotion Review, 3*(4), 364-370. https://doi.org/10.1177/1754073911410740

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior, 1*(1), 56-75. https://doi.org/10.1007/BF01115465

Field, A. (2013). *Discovering statistics using IBM SPSS* (5th ed.). Sage Publications.

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., Baron-Cohen, S., & Golan, O. (2016). Basic and complex emotion recognition in children with autism: cross-cultural findings. *Molecular Autism, 7*(1), 1-11. https://doi.org/10.1186/s13229-016-0113-9

Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The 'reading the mind in films' task [child version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders, 38*(8), 1534-1541. https://doi.org/10.1007/s10803-007-0533-7

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 36*(2), 169-183. https://doi.org/10.1007/s10803-005-0057-y

Golan, O., Sinai-Gavrilov, Y., & Baron-Cohen, S. (2015). The Cambridge Mindreading Face-Voice Battery for Children (CAM-C): Complex emotion recognition in children with and without autism spectrum conditions. *Molecular Autism, 6*(1), 1-9. https://doi.org/10.1186/s13229-015-0018-z

Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). McGraw-Hill.

Hallerbäck, M. U., Lugnegård, T., Hjärthag, F., & Gillberg, C. (2009). The Reading the Mind in the Eyes Test: Test-retest reliability of a Swedish version. *Cognitive Neuropsychiatry, 14*(2), 127-143. https://doi.org/10.1080/13546800902901518

Iglesias, J., Loeches, A., & Serrano, J. (1989). Expresión facial y reconocimiento de emociones en lactantes. *Infancia y Aprendizaje, 12*(48), 93-113. https://doi.org/10.1080/02103702.1989.10822251

International Test Commission. (2014). *ITC statement on the use of tests and other assessment instruments for research purposes.*

Izard, C. E. (1977). Differential emotions theory. In C. E. Izard, *Human emotions* (pp. 43-66). Springer.

Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition & Emotion, 13*(5), 505-521. https://doi.org/10.1080/026999399379168

Kennedy, D. P., & Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends in Cognitive Sciences, 16*(11), 559-572. https://doi.org/10.1016/j.tics.2012.09.006

Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. *Sensors, 18*(2), 401. https://doi.org/10.3390/s18020401

Kosonogov, V., & Titova, A. (2019). Recognition of all basic emotions varies in accuracy and reaction time: A new verbal method of measurement. *International Journal of Psychology, 54*(5), 582-588. https://doi.org/10.1002/ijop.12512

Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review, 5*(1), 41-46. https://doi.org/10.1177/1754073912451349

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*(1), 155-184. https://doi.org/10.1146/annurev.neuro.23.1.155

Leiva, S. (2017). Validación de una batería para evaluar el reconocimiento de emociones a través del rostro y del cuerpo utilizando estímulos dinámicos. *Revista Argentina de Ciencias del Comportamiento (RACC), 9*(3), 60-81. https://doi.org/10.32348/1852.4206.v9.n3.17186

Lieberman, M. (2010). Social cognitive neuroscience. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 446-496). McGraw-Hill.

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología/Annals of Psychology, 30*(3), 1151-1169. https://doi.org/10.6018/analesps.30.3.199361

Martínez Arias, M. R., Hernández Lloreda, M. V., & Hernández Lloreda, M. J. (2014). *Psicometría.* Alianza Editorial.

Matsumoto, D., & Ekman, P. (1988). *Japanese and Caucasian Facial Expressions of Emotion and Neutral Faces* (JACFEE and JACNeuF). Human Interaction Laboratory, University of California, San Francisco, 401 Parnassus Avenue, San Francisco, CA, 94143.

McDonald, S., & Saunders, J. C. (2005). Differential impairment in recognition of emotion across different media in people with severe traumatic brain injury. *Journal of the International Neuropsychological Society, 11*(4), 392-399. https://doi.org/10.1017S1355617705050447

Medrano, L. A., Moretti, L., Ortiz, Á., & Pereno, G. (2013). Validación del cuestionario de regulación emocional cognitiva en universitarios de Córdoba, Argentina. *Psykhe, 22*(1), 83-96. https://doi.org/10.18682/pd.v21i1.3881

Morales, H. L., Agulla, L., Zabaletta, V., Vivas, L., & López, M. (2017). Rostros, gestos y emociones: procesamiento diferencial de las expresiones faciales emocionales en población infanto-juvenil según el sexo. *Revista Argentina de Ciencias del Comportamiento, 9*(3), 31-43. https://doi.org/10.32348/1852.4206.v9.n3.18521

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

Recio, G., Schacht, A., & Sommer, W. (2013). Classification of dynamic facial expressions of emotion presented briefly. *Cognition & Emotion, 27*(8), 1486-1494. https://doi.org/10.1080/02699931.2013.794128

Richaud, M. C., Lemos, V. N., Mesurado, B., & Oros, L. (2017). Construct validity and reliability of a new Spanish empathy questionnaire for children and early adolescents. *Frontiers in Psychology, 8,* Article 979. https://doi.org/10.3389/fpsyg.2017.00979

Rodgers, J. D., Lopata, C., Booth, A. J., Thomeer, M. L., Donnelly, J. P., Rajnisz, C. J., Wood, J. T., Lodi-Smith, J., & Kozlowski, K. F. (2021). Psychometric properties of the Cambridge-Mindreading Face-Voice Battery for Children in children with ASD. *Autism Research, 14*(9), 1965-1974. https://doi.org/10.1002/aur.2546

Schweizer, K., Ren, X., & Zeller, F. (2019). On modeling the ceiling effect observed in cognitive data in the framework of confirmatory factor analysis. *Psychological Test and Assessment Modeling, 61*(3), 333-353.

Soto, C. M., & Segovia, J. L. (2009). Intervalos de confianza asimétricos para el índice la validez de contenido: un programa Visual Basic para la V de Aiken. *Anales de Psicología/Annals of Psychology, 25*(1), 169-171. https://doi.org/10.6018/analesps

Suzuki, A., Hoshino, T., & Shigemasu (2006). Measuring individual differences in sensitivities to basic emotions in faces. *Cognition, 99*(3), 327-353. https://doi.org/10.1016/j.cognition.2005.04.003

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson. https://books.google.com/ books?id=2KJIngEACAAJ

Thomeer, M. L., Smith, R. A., Lopata, C., Volker, M. A., Lipinski, A. M., Rodgers, J. D., McDonald, C. A., & Lee, G. K. (2015). Randomized controlled trial of mind reading and in vivo rehearsal for high-functioning children with ASD. *Journal of Autism and Developmental Disorders, 45*(7), 2115-2127. https://doi.org/10.1007/s10803-015-2374-0

Vaci, N., Gula, B., & Bilalić, M. (2014). Restricting range restricts conclusions. *Frontiers in Psychology, 5,* Article 569. https://doi.org/10.3389/fpsyg.2014.00569

Vatandoust, L., & Hasanzadeh, R. (2018). The study of the emotion recognition and the cognitive Failures of children with developmental coordination disorder. *Iranian Rehabilitation Journal, 16*(2), 121-130. https://doi.org/10.32598/irj.16.2.121

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry, 18*(4), 326-354. https://doi.org/10.1080/13546805.2012.721728

Walle, E. A., Lopez, L. D., & Merced, C. A. (2020). Emotion recognition and understanding in infancy and early childhood. *Encyclopedia of Infant and Early Childhood Development,* 537-545. https://doi.org/10.1016/B978-0-12-809324-5.23567-0

Wilson-Mendenhall, C. D., Barrett, L. F., & Barsalou, L. W. (2013). Neural evidence that human emotions share core affective properties. *Psychological Science, 24*(6), 947-956. https://doi.org/10.1177/0956797612464242

Wingenbach, T. S., Ashwin, C., & Brosnan, M. (2018). Sex differences in facial emotion recognition across varying expression intensity levels from videos. *PLoS one, 13*(1), Article e0190634. https://doi.org/10.1371/journal.pone.0190634

Zupan, B., & Neumann, D. (2016). Exploring the use of isolated expressions and film clips to evaluate emotion recognition by people with traumatic brain injury. *JoVE (Journal of Visualized Experiments),* (111), Article e53774. https://doi.org/10.3791/53774