UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Mejia, M. A., Cardoso, A., Lozoya, V., & Bobes, A. (2026). Adaptation of the 20-Item Prosopagnosia Index for the Screening of Developmental Prosopagnosia in Mexico. *Collabra: Psychology*, *12*(1). https://doi.org/10.1525/collabra.163055

Clinical Psychology

Adaptation of the 20-Item Prosopagnosia Index for the Screening of Developmental Prosopagnosia in Mexico

Manuel Alejandro Mejia[1], Agustin Cardoso[1][a], Vania Lozoya[1], Antonieta Bobes[2]

[1] CETYS University, Baja California, Mexico, [2] Cognitive Neurosciences, Cuban Center for Neurosciences, Havana, Cuba

Keywords: developmental prosopagnosia, face recognition, PI20, cross-cultural adaptation, psychometric validity, reliability, Mexican population, self-report, Cambridge Face Memory Test, Cambridge Car Memory Test, neuropsychological assessment

https://doi.org/10.1525/collabra.163055

Collabra: Psychology

Vol. 12, Issue 1, 2026

Developmental prosopagnosia (DP) is a lifelong condition characterized by difficulties in recognizing faces despite intact vision and intelligence. The 20-Item Prosopagnosia Index (PI20) is a widely used self-report instrument for screening face recognition deficits. However, no standardized Spanish version of the PI20 existed prior to this study, limiting its applicability in Mexican populations and other Spanish-speaking contexts where no validated version exists. This study aims to adapt and validate the PI20 for use in Mexican Spanish, supporting its linguistic and contextual appropriateness for use in the Mexican population. Study 1 involved a rigorous translation and cultural adaptation process, including forward translation by multiple independent translators, expert panel review, and cognitive interviews with bilingual Mexican participants. Qualitative analysis of these interviews identified semantic and conceptual issues in several items, which were subsequently revised. A pilot test with 15 participants confirmed the clarity and feasibility of the adapted version, supporting its content validity. Study 2 evaluated the psychometric properties of the adapted PI20 in a sample of 333 adults. The results showed high internal consistency (McDonald's omega and Cronbach's alpha = .84) and good test–retest reliability ($ICC_{2,1}$= .81). Confirmatory factor analysis supported a unidimensional structure consistent with the original scale, while exploratory analysis revealed two correlated components that did not undermine the overall model. The adapted PI20 showed a moderate negative correlation with the Cambridge Face Memory Test ($r(316) = -.229$, $p < .001$), supporting convergent validity, and no significant association with the Cambridge Car Memory Test ($r(313) = -.106$, $p = .061$), supporting discriminant validity. These findings suggest that the Mexican Spanish version of the PI20 is a reliable and valid instrument for assessing self-reported facial recognition difficulties in Spanish-speaking Mexican adults. This adaptation represents an important step toward the availability of linguistically and contextually appropriate tools for research and screening purposes.

Introduction

Prosopagnosia, or "face blindness," is a neurological disorder characterized by difficulty or inability to recognize familiar faces. This condition can be present from birth, known as developmental prosopagnosia (DP), or acquired as a consequence of brain injury (Garrido et al., 2008). DP can have significant repercussions on the social and interpersonal lives of those affected, often leading to difficulties in everyday social interactions, maintaining interpersonal relationships, and navigating common social situations (Corrow et al., 2016). The prevalence of DP has been estimated to range between 0.93% and 3.08% in samples of adults from the United States (DeGutis et al., 2023), suggesting that face recognition difficulties are not uncommon in the general population. Consistent with this observation, face recognition ability is thought to vary continuously across the population, with developmental prosopagnosia representing the lower end of this distribution rather than a strictly categorical clinical condition (Duchaine & Nakayama, 2006; Shah et al., 2015). However, these findings are not globally representative, and there are currently no prevalence data available for the Mexican population.

In addition to its impact on everyday social functioning, DP provides an important window into the cognitive and

a Corresponding Author: agustin.cardoso@cetys.edu.mx

neural mechanisms that support face perception more broadly. Important theoretical questions remain regarding the basic mechanisms that support face recognition and whether these mechanisms are specialized for faces. Research on DP has provided convergent evidence that face perception relies on the global integration of facial features rather than on the isolated analysis of individual components (Corrow et al., 2016). Although some individuals with DP experience additional difficulties with object or body recognition (Brewer et al., 2018; Stumps et al., 2020), evidence suggests that the core deficit primarily affects facial identity processing, while other perceptual abilities, such as color perception, often remain intact (Smith & Susilo, 2021). Consistent with these behavioral findings, neuroimaging studies suggest that DP is specifically associated with cortical thinning in key areas of the face recognition network, particularly the fusiform face area (FFA) and the occipital face area (OFA), which are involved in the holistic processing of facial information (Anzellotti & Caramazza, 2014; Towler et al., 2017). Recent research also suggests that DP may frequently co-occur with other neurodevelopmental conditions. For example, Maw et al. (2025) reported that up to 56% of individuals with probable developmental coordination disorder met recommended criteria for prosopagnosia, highlighting a substantial overlap between these conditions. Taken together, the variability in face recognition ability and its overlap with other neurodevelopmental conditions underscore the importance of accurately assessing individual differences in face recognition abilities.

Proper assessment of developmental prosopagnosia is essential for understanding individual differences in face recognition ability and for identifying individuals who experience persistent difficulties in recognizing faces. Current approaches typically rely on a combination of objective behavioral tasks and self-report measures. Objective tests commonly include measures of face perception, such as the Cambridge Face Perception Test and the Benton Facial Recognition Test (Benton et al., 1994), as well as tests of memory for unfamiliar faces, such as the *Cambridge Face Memory Test* designed by Duchaine & Nakayama (2006). In addition, tests of memory for famous faces are frequently used to evaluate real-world face recognition abilities. Measures such as the Cambridge Car Memory Test (CCMT) are also often included to verify that difficulties are specific to faces rather than reflecting broader perceptual or memory impairments.

Alongside these objective measures, self-report instruments play an important role in identifying lifelong difficulties with face recognition that may not always be fully captured by laboratory tasks. One of the most widely used instruments for this purpose is the 20-Item Prosopagnosia Index (PI20), which has shown meaningful associations with objective measures of face recognition and has proven effective in identifying symptoms associated with prosopagnosia (Shah et al., 2015).

The PI20 has been adapted into several languages—such as Japanese, French, Danish, Chinese, Italian, and Portuguese (Nakashima et al., 2020; Nigrou et al., 2024; Nørkær et al., 2023; Sun et al., 2021; Tagliente et al., 2023; Ventura et al., 2018)—, but there is not a standardized Spanish version. This shortcoming limits its application in Mexican Spanish-speaking contexts, where cultural and linguistic norms may influence face perception and evaluation. In the Spanish-speaking world, and specifically in Mexico, there are few studies on prosopagnosia and limited adapted diagnostic tools, highlighting the need for validated instruments in these contexts.

The present study aims to examine the psychometric properties of the Mexican Spanish version of the PI20 with a mixed-methods design to provide a reliable and culturally appropriate instrument for the evaluation of developmental prosopagnosia, in accordance with international standards (Arizpe et al., 2019; Hambleton & Lee, 2013). This process involves a first study that aimed at translating the scale from British English into Spanish, comparing the translated version with the original, and evaluating its semantic and conceptual validity through expert feedback, cognitive interviews, and a pilot test. The objective was to ensure that the instrument is clear, culturally relevant, and reliable for assessing self-reported face recognition difficulties in the Mexican Spanish-speaking populations. Finally, a second study entailed evaluating the psychometric properties of the Spanish version of the PI20 in a larger sample.

Validation Study 1: Adapted Version of the PI20, and Evidence of Content Validity

The first study aimed to construct an initial Spanish version of the PI20 by following current guidelines of test adaptation, with the objective of producing a culturally appropriate version. The translation was carried out by independent translators, evidence to support content validity was produced from interviews with a variety of experts as well as cognitive interviews.

Methods

Participants

Three different sets of participants were recruited. The first group was composed of three professional translators with Spanish as their native language and professional fluency in English. These translators worked independently from each other to ensure that the initial translation would reflect different perspectives and reduce potential bias.

The second group of participants was selected to take part in cognitive interviews. This group consisted of five adult individuals, two men and three women, between 24 and 47 years old, who were bilingual (L1: Spanish, L2: English), and residents of Tijuana, Mexico (see Table 1). Tijuana is a border city with distinctive sociocultural characteristics due to its proximity to the United States, which results in frequent bilingual and bicultural dynamics among its population. In addition, the city has a highly diverse population due to internal migration within Mexico. According to national census data, approximately 49% of the population of Tijuana was born outside the state of Baja California, and around 7% of residents arrived in the city

Table 1. Sociodemographic characteristics of participants involved in study 1 (n = 20).

Variables	Cognitive interview participants (n = 5)	Pilot participants (n = 15)
Age mean (SD)	28.58	37 (SD: 7.9)
Age range	24 - 47	23 - 49
Gender		
Women	3	10
Men	2	5
Education		
High School	2	10
Undergraduate	3	5
Language profile	Spanish (L1), English (L2) bilinguals	Spanish (L1). L2 not recorded

Note. Cognitive interview participants were bilingual adults recruited to evaluate item clarity and interpretation during the cultural adaptation process. Pilot participants completed the preliminary Spanish version of the PI20 to evaluate clarity of instructions, response time, and preliminary functioning of the instrument.

within the last five years (INE, 2020). As bilingual participants, they were also able to evaluate the semantic equivalence between the original English items and their Spanish translations during the interview process. Recruitment for this group was carried out through digital diffusion strategies, such as online announcements and social media posts. Their role was to provide feedback on comprehension and interpretation of the items.

Finally, the third group was recruited to complete a pilot version of the Spanish PI20. For this stage, 15 additional participants between 23 and 49 years of age were recruited at a private university in Tijuana, Mexico to assess the clarity of instructions, the overall response time, and the internal consistency of the preliminary version of the instrument (see Table 1). The data obtained with the third group guided final adjustments before proceeding to the larger scale psychometric validation in Study 2.

Instruments

The primary instrument is the PI20 (Shah et al., 2015), a 20-item questionnaire designed to identify difficulties in facial recognition through a Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). The total score, which ranges between 20 and 100, is expected to reflect the degree of perceived difficulty, with scores above 65 being indicative of possible prosopagnosia. Previous studies have shown that the PI20 has high internal consistency (Cronbach’s alpha between 0.82 and 0.88; McDonald’s omega ≈ 0.81) and evidence of convergent validity—correlating with instruments such as the Cambridge Face Memory Test (CFMT)—as well as discriminant validity, demonstrated by the lack of correlation with tests assessing object recognition, such as the Cambridge Car Memory Test (CCMT) (Nakashima et al., 2020; Nigrou et al., 2024; Nørkær et al., 2023; Shah et al., 2015; Sun et al., 2021; Ventura et al., 2018).

Additionally, to assess cultural equivalence of the PI20 after translation into Spanish, we used a short version of the Hambleton & Zenisky (2010) questionnaire. This questionnaire used 19 items grouped into four categories: general criteria, item format, grammar and phrasing, and culture. Six items from the original form were excluded because they were only relevant for educational tests (item 9) and if passages are present (items 16-20). Participants judged how accurate the translation was by answering “Yes”, “No”, “Unsure” or “Not relevant” after comparing each item in its original English version with the Spanish adapted version.

Procedure

The adaptation process of the PI20 to Mexican Spanish followed the recommendations established in the *International Test Commission Guidelines for Translating and Adapting Tests* (International Test Commission, 2017). A cross-sectional, descriptive mixed-methods design was employed, integrating quantitative data obtained from the scores of the adapted PI20 and qualitative data collected through cognitive interviews, following recommendations for cognitive interviewing procedures in questionnaire development (as per the recommendation by Sexton et al., 2023). This approach allowed the identification of potential misunderstandings and supported the evaluation of conceptual equivalence between the original and adapted versions. In summary, the translation process followed adaptation guidelines by relying on multiple independent forward translations, expert committee review, reconciliation of translation discrepancies, and cognitive interviews with members of the target population to ensure semantic and conceptual equivalence of the adapted items.

First, three independent bilingual translators produced separate forward translations of the original instrument. These preliminary translations were subsequently compared and discussed in a consensus meeting with an expert panel composed of four specialists. The panel included a researcher with expertise in psychometrics and neuropsychology (first author), a researcher with extensive experience in prosopagnosia and face recognition research (last author), and two experts with recognized experience in Mexican cultural contexts and training in psychometrics (listed in the acknowledgements). During this stage, the translations were systematically compared to evaluate semantic, conceptual, and cultural equivalence. A reconciled preliminary Spanish version of the PI20 was then produced and presented back to the translators, who provided additional feedback and suggestions. Both the expert panel and the subsequent cognitive interviews contributed to refining items that initially lacked consensus among translators, particularly items 8, 18, and 20.

Cognitive interviews were subsequently conducted using the preliminary Spanish version of the PI20 to evaluate the clarity, interpretability, and cultural relevance of the translated items. This phase aimed to identify potential ambiguities or misunderstandings that could arise when respondents interpreted the questionnaire. Participants were asked to explain their understanding of each item and describe any difficulties they experienced while answering the questionnaire. The interviews followed a semi-structured format that included paraphrasing and verbal probing techniques commonly used in cognitive interviewing methodologies to explore how respondents interpret survey questions and response options.

Additionally, five bilingual participants (L1: Spanish, L2: English), two men and three women (22-40 years old), compared the English and Spanish versions to evaluate whether the translated wording preserved the meaning and intent of the original statements, providing additional evidence of linguistic equivalence. The method for this comparison was proposed by Hambleton & Zenisky (2010) specifically to strengthen the judgements in adaptations processed following the ITC (2017) Guidelines for Test Adaptation. This method uses a standardized form that asks specific questions about the most usual adaptation problems in cross-cultural research, and has been widely recommended for test adaptations (Nguyen et al., 2024; Reynolds & Suzuki, 2012).

Subsequently, a pilot study with 15 participants was conducted to evaluate the clarity of the instructions, response time, and to collect suggestions from participants. The results of this pilot phase informed the final adjustments to the instrument before proceeding to the large-scale psychometric validation conducted in Study 2.

Small samples are commonly used in the preliminary phases of cross-cultural test adaptation because the primary objective of these stages is to evaluate the clarity, interpretability, and semantic equivalence of translated items rather than to conduct quantitative psychometric analyses. Methodological literature on the adaptation of self-report instruments recommends the use of relatively small pilot samples when procedures such as cognitive interviews or pretesting are employed to identify wording problems, ambiguities, or interpretation difficulties before large-scale validation studies (Beaton et al., 2000; International Test Commission, 2017; Sexton et al., 2023).

Ethical Considerations

The study was conducted in accordance with the principles of the declaration of Helsinki ("World Medical Association Declaration of Helsinki," 2013). All participants signed written informed consent, and the research protocol was approved by the authors institutional ethics committee (approval number #CEI-154), ensuring confidentiality and the proper use of the collected data.

Results

First, the work done by the three independent translators and the subsequent experts committee produced a draft version of the PI20 in Spanish. The translators reached an agreement in all items, except items 8, 18, and 20. The result from this first step was the draft used for the cognitive interviews. The analyses of the cognitive interviews revealed the need to adjust certain terms and phrases to improve comprehension in the Mexican context. These findings are summarized in Table 2.

The cognitive interviews helped reach a consensus on items 8, 18, and 20, which had not been achieved during the initial translation phase between the independent translators. The results of the cognitive interviews resolved ambiguities and the translations were improved, ensuring their suitability for the Mexican context and that the items were interpreted in a way that accurately reflected experiences of face recognition difficulties among Mexican participants. This comparison indicated that the Spanish wording adequately retained the conceptual meaning of the original PI20 items. These data can be seen in Table 3, and the items in Spanish are shown in Table S1.

Finally, the PI20 version was assessed by five bilingual participants who judged if the translated version preserved the meaning and intent of the original statements in four categories: general characteristics, item format, grammar and phrasing, and culture. Average expert agreement across items in these categories ranged from 82% to 99% (M = 91%). All items showed agreement above 83% in the category for cultural equivalence (M = 96%). Most discrepancies were related to length (item 11) and linguistic complexity (items 3, 4, and 11) rather than conceptual or cultural differences (see a summary for the 20 items in Table S2).

Conclusion of Study 1

A Spanish version of the PI20 was produced following international guidelines, including independent translators, expert consensus, cognitive interviews, and pilot testing. The strength of this process lies in the use of cognitive interviews as a systematic method of analysis, which allowed for the detailed exploration of how participants understood, interpreted, and responded to each item. This qualitative approach ensured that potential ambiguities or cultural mismatches were identified and corrected in real time. The results provided preliminary support for the linguistic and contextual appropriateness of the Mexican Spanish version of the PI20, as participants provided concrete feedback that led to the resolution of problematic items—particularly items 8, 18, and 20—thereby improving the clarity, contextual relevance, and semantic equivalence of the scale. Taken together, these procedures provide preliminary evidence supporting content validity and a strong foundation for Study 2 to estimate the psychometric properties of the instrument.

Table 2. Findings from Cognitive Interviews on the Spanish Adaptation of the PI20.

Items	Findings	Actions
1 to 7	The first seven items were well understood by the participants, especially those related to self-assessment of the ability to recognize familiar faces and memorize faces, which presented little difficulty. The wording was considered clear and accessible.	Although they were generally well understood, some minor wording details were reviewed to optimize fluency and ensure there were no ambiguities.
8	Most participants adequately understood the intention of the item as the ability to generate a mental image of a person's face. However, a few isolated cases interpreted “imaginarme” as recalling past experiences, which evidenced the need to retain clear wording that would orient toward visual imagery.	The verb “imaginarme” was retained in the final wording to remain faithful to the original intention of the instrument, which seeks to assess the ability to mentally visualize faces and not the memory of past events or personal associations.
9 to 17	These items were mostly well received by the participants, with some minor observations about the wording. However, participants did not report significant comprehension issues.	The items were considered to be aligned with the participants' daily experiences, and no major adjustments were required for most of them.
18	In the cognitive interviews, participants understood the item in general, although with different interpretations of the term “confundir” (faces, names or identities). Responses showed that the formulation was functional in a family context and reflected real experiences, especially in large gatherings.	The wording “A veces confundo a los miembros de mi familia cuando nos reunimos” was retained, as it was adequately understood by the majority and retains the intent of the original construct.
19	Participants understood the item and related it to real experiences, such as seeing comparisons of celebrities before and after fame. They mentioned concrete examples (e.g. celebrities that have undergone surgeries) and, although some pointed out that certain changes make recognition difficult, in general they considered the statement clear and relevant.	The version: “Me resulta fácil reconocer a celebridades en imágenes de ‘antes que fueran famosos’ aunque hayan cambiado considerablemente” was retained, as it was correctly understood and adequately reflects the construct assessed.
20	Participants understood the item and related it to situations in which they see people they know outside their usual context. Although most of them indicated that they do not have difficulties in these cases, they recognized the scenario as a possible and clear experience.	The wording “Se me complica reconocer a personas conocidas cuando las veo fuera de situaciones o lugares habituales (ej., al encontrarme inesperadamente a un compañero de trabajo al hacer compras)” was maintained, as it was adequately understood and reflects the construct assessed.

Table 3. Consensus on Items 8, 18, and 20 of the Spanish Adaptation of the PI20.

Items	Findings	Actions
8: "I find it easy to picture individual faces in a crowd" Proposed translation: “Se me facilita imaginarme el rostro de alguien en medio de una multitud.”	During translation, no consensus was reached on the use of “alguien” as opposed to terms such as “individuos”. Some translators proposed more specific alternatives, although without clear agreement.	The item was reformulated as: “Se me facilita imaginarme el rostro de alguien en mi mente”, as this version was more natural, general and understandable for the participants during the cognitive interviews.
18: "At family gatherings, I sometimes confuse members of my family with others" Proposed translation: “A veces confundo a los miembros de mi familia con otras personas en reuniones familiares.”	Although the sentence was considered clear by the translators, in the cognitive interviews different interpretations of the verb “confundir” (faces, names or identities) were observed, so it was suggested to simplify the structure.	The final version was adjusted to: “A veces confundo a los miembros de mi familia cuando nos reunimos”, as it was considered clearer and more in line with Mexican Spanish usage, maintaining the intention of the original item.
20: "It is hard to recognize familiar people when they are out of context" Proposed translation: “Se me complica reconocer a personas conocidas fuera de su contexto habitual.”	Some translators expressed doubts about the clarity of the phrase “fuera de contexto” and suggested specifying the setting. Cognitive interviews confirmed that the difficulty arose in the interpretation of unexpected or unusual settings.	A contextual example was incorporated to improve clarity: “Se me complica reconocer a personas conocidas cuando las veo fuera de situaciones o lugares habituales (ej. al encontrarme inesperadamente a un compañero de trabajo al hacer compras)” which facilitated their understanding during cognitive interviews.

Validation Study 2: Reliability and Evidence of Validity

The purpose of the second study was to gather evidence of reliability and construct validity of the Spanish version of the PI20. For this objective, a cross-sectional correlational study was conducted with a larger sample and the inclusion of additional face and object recognition tests. The study tested for evidence of internal structure, convergent validity with face perception and face memory, and discriminant validity with object memory.

Methods

Participants

A total of 333 participants completed the adapted version of PI20 from Study 1. These participants were recruited primarily from universities and community networks in the city of Tijuana, Mexico, including the 15 participants from the pilot in Study 1. The additional 318 participants recruited for Study 2 also completed a face recognition battery described below. Two participants did not complete the battery, but their data was included for tasks completed fully (one task for one participant, and two tasks for the other). The sample included 127 women (38.14%) and 206 men (61.86%), between 18 to 65 years old (M = 28.58, SD = 11.8). Participants were recruited through non-probabilistic convenience sampling, with efforts made to ensure diversity in terms of age, educational background and socioeconomic status (see Table 4). All participants were required to be able to read and complete the questionnaire independently as part of the inclusion criteria.

The sample size was estimated from power analyses. For the correlations between tests, power analyses run with the *pwr* package in R (Champely, 2020), in RStudio (Posit team, 2023), indicated that a sample size of at least 164 would have 90% power to detect effects of r > .25 with an alpha level of 0.05. For the evidence of internal structure, power analysis for a confirmatory factor analysis (CFA) indicated that a sample size of 147 would have 90% power to detect a root mean square error of approximation (RMSEA) of 0.05 for a model of a single factor with 20 items (using calculator by Arifin, 2026; that uses calculations from Kim, 2005).

For internal consistency and factor structure analyses, data from all 333 participants with complete responses across all PI20 items was used. Additionally, a subsample of 31 participants completed a second session after approximately six weeks to assess test-retest reliability. Post-hoc sensitivity analyses using the R package ICC.Sample.Size indicated that the largest intraclass correlation that could be rejected as null-hypothesis was 0.58 (80% power, with an alpha of 0.05).

Table 4. Sociodemographic characteristics of participants who completed the PI20 and the face recognition battery (Study 2, n = 333).

Variables	M / F	SD / %
Age in years	28.58	11.8
Gender		
Women	127	38.14
Men	206	61.86
Education		
No education	2	0.6
Primary	4	1.2
Secondary	41	12.31
High School	206	61.86
Undergraduate	75	22.52
Graduate studies	5	1.5

Note. M = Mean, F = Frequency, SD = Standard Deviation
The category "no education" refers to participants who did not complete formal schooling and does not imply illiteracy. All participants were able to read and complete the questionnaire independently.

Instruments

20-Item Prosopagnosia Index (PI20)

The primary instrument used in this study was the adapted 20-Item Prosopagnosia Index (PI20; Shah et al., 2015) from Study 1, a self-report questionnaire designed to identify difficulties in face recognition. The scale consists of 20 items rated on a five-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). Total scores range from 20 to 100, with higher scores indicating greater self-reported difficulties in recognizing faces. Previous studies have reported high internal consistency for the PI20 (Cronbach's alpha ranging between .82 and .88) as well as evidence of convergent validity with objective measures of face recognition, such as the Cambridge Face Memory Test, and discriminant validity with object recognition tasks such as the Cambridge Car Memory Test (Nakashima et al., 2020; Nørkær et al., 2023; Shah et al., 2015; Sun et al., 2021; Ventura et al., 2018).

Cambridge Face Memory Test (CFMT)

The Cambridge Face Memory Test is a widely used behavioral measure designed to assess memory for unfamiliar faces (Duchaine & Nakayama, 2006). In this task, participants first learn a set of unfamiliar faces and are subsequently tested on their ability to recognize them across different viewpoints and visual conditions. The CFMT has demonstrated strong reliability and validity and is considered a standard objective measure of face memory in research on face recognition and developmental prosopagnosia.

Cambridge Face Perception Test (CFPT)

The Cambridge Face Perception Test evaluates face perception abilities independently of memory processes (Duchaine et al., 2007). In this task, participants are required to sort faces according to their similarity to a target face, allowing the evaluation of perceptual discrimination of facial identity. The CFPT is widely used in research on face processing and developmental prosopagnosia.

Cambridge Car Memory Test (CCMT)

The Cambridge Car Memory Test is an object memory task designed to parallel the structure of the CFMT while assessing memory for non-face objects (Dennett et al., 2012). The task uses images of cars instead of faces and allows researchers to determine whether performance differences are specific to faces or reflect broader visual memory abilities. For this reason, the CCMT is commonly used as a control task in studies of face recognition.

For the CFMT, CFPT, and CCMT, the task instructions were translated into Spanish by the authors of this study while maintaining the original stimuli and experimental structure.

Procedure

The Mexican Spanish version of the PI20 obtained from Study 1 was administered on a large scale under controlled conditions alongside the standardized battery of face processing tests described above. Participants completed the PI20 together with the Cambridge Face Memory Test (CFMT), the Cambridge Face Perception Test (CFPT), and the Cambridge Car Memory Test (CCMT), in that order.

The tasks were implemented in PsychoPy (Peirce et al., 2019) and administered online through Pavlovia (https://pavlovia.org/). The CFMT and CFPT were used to assess convergent validity of the PI20, whereas the CCMT was included to assess discriminant validity. This methodological approach allowed a direct comparison between self-reported face recognition difficulties measured by the PI20 and objective performance in face and object recognition tasks.

The study was conducted in full compliance with the ethical procedures established in Study 1. Data analyses were performed using JASP (Version 0.95.1).

Results

Descriptive Results of the Face Battery

Table 5 presents the descriptive statistics for PI20 and the face processing battery. The average total score on the PI20 was 43.68 (SD = 9.36), with higher scores indicating higher self-reported difficulty in face recognition. In the Cambridge Face Memory Test (CFMT), the mean score was 54.41 correct trials (SD = 10.26; N = 318; maximum score = 72). The Cambridge Car Memory Test (CCMT) had an average of 45.69 correct trials (SD = 9.44; N = 315; maximum score = 72), while the Cambridge Face Perception Test (CFPT), which is scored by total errors for upright faces, showed a mean of 48.99 (SD = 20.99; N = 316).

Reliability of the PI20

The Spanish version of PI20 demonstrated strong internal consistency and temporal reliability. The McDonald's Omega and Cronbach's Alpha coefficients for the full PI20 scale were both 0.84. Using the method proposed by Bonett and Wright (2015), the 95% confidence intervals of the Spanish version were estimated to range from 0.82 to 0.87 for Omega and from 0.81 to 0.86 for Alpha, indicating robustness of the results. The intraclass correlation coefficient (ICC2,1), used to assess test-retest stability, was 0.81(95% CI [0.762,0.844], n=31), further supporting the scale's temporal robustness.

Test-retest reliability was evaluated using the intraclass correlation coefficient (ICC). The ICC is widely recommended for reliability studies because it allows the estimation of agreement between repeated measurements obtained on the same scale and provides an indicator of score stability across measurement occasions (Koo & Li, 2016). In the present study, the ICC(2,1) model was used, corresponding to a two-way random effects model, which is appropriate when repeated measurements are considered interchangeable and when the objective is to evaluate the overall stability of the instrument beyond a specific administration (Koo & Li, 2016; Liljequist et al., 2019).

The subsample used for the test–retest analysis included 31 participants who completed the PI20 again after a six-week interval. This number is smaller than the total sample because participation in the second measurement depended on voluntary follow-up participation. Therefore, the results should be interpreted as initial evidence of temporal stability, and future studies with larger samples will allow these estimates to be confirmed and refined. In reliability studies, test–retest analyses are commonly conducted on subsamples of the original participants because repeated participation across time often results in lower response rates. Methodological literature indicates that reliability estimates can still be informative with moderate sample sizes, particularly when the obtained coefficients are relatively high (Shoukri et al., 2004).

Comparative estimates of internal consistency for the face processing battery also indicated high reliability. The CFMT showed the highest internal consistency (ω = 0.91; α = 0.90), while the CCMT and CFPT demonstrated solid reliability as well (ω = 0.84 and 0.83; α = 0.84 and 0.82, respectively). The test-retest values for the face processing battery presented in Table 5 exceed the .5 threshold for all tests, values ranging from 0.66 (CCMT) to 0.75 (CFPT), indicating moderate to good intraclass correlation (Koo & Li, 2016).

Together, these results provide preliminary support for the internal structure of the PI20 and suggest that the adaptation retains adequate psychometric properties when compared to established face and object recognition instruments, as shown in Table 6. Furthermore, as argued by Stensen & Lydersen (2022), McDonald's omega may offer a more appropriate measure of internal consistency than Cronbach's alpha, particularly when the assumption of tau-equivalence is questionable.

Item-level analyses of PI20 supported the internal reliability of the adapted scale. McDonald's Omega coefficients ranged from 0.826 to 0.858, and Cronbach's Alpha values ranged from 0.815 to 0.851 across the 20 items, indicating consistent item performance. Although the removal of items 8 and 19 produced marginal increases in reliability estimates, the improvements were not substantial enough to justify their exclusion. These results suggest that all

Table 5. Descriptive statistics of the Spanish version of the PI20 and the Cambridge battery (CFMT, face memory; CCMT, car memory; CFPT, face perception) in Study 2.

Test	Mean	SD	Min	Max	N
CFMT Total	54.41	10.26	29	72	318
CCMT Total	45.69	9.44	23	72	315
CFPT Total errors	48.99	20.99	10	114	316
PI20 Total	43.68	9.36	21	70	333
PI20 Item totals					
PI20 Item 01	2.23	0.93	1	5	333
PI20 Item 02	2.29	0.99	1	5	333
PI20 Item 03	3.45	1.00	1	5	333
PI20 Item 04	2.05	0.97	1	5	333
PI20 Item 05	1.59	0.71	1	4	333
PI20 Item 06	1.91	0.85	1	5	333
PI20 Item 07	1.81	0.93	1	5	333
PI20 Item 08	2.69	1.17	1	5	333
PI20 Item 09	3.11	0.99	1	5	333
PI20 Item 10	2.14	0.87	1	5	333
PI20 Item 11	1.79	0.91	1	5	333
PI20 Item 12	2.13	0.97	1	5	333
PI20 Item 13	1.96	1.02	1	5	333
PI20 Item 14	1.88	0.92	1	5	333
PI20 Item 15	1.69	0.87	1	5	333
PI20 Item 16	1.86	0.95	1	5	333
PI20 Item 17	2.80	1.14	1	5	333
PI20 Item 18	1.39	0.63	1	5	333
PI20 Item 19	2.85	1.04	1	5	333
PI20 Item 20	2.11	0.97	1	5	333

Table 6. Reliability indices for the Spanish version of the PI20 (CFMT, face memory; CCMT, car memory; CFPT, face perception; Study 2).

Test	Coefficient	Estimate	Lower	Upper	Std. Error
PI20	McDonald's ω	0.84	0.82	0.87	0.01
	Cronbach's α	0.84	0.81	0.86	0.01
	ICC(2,1)	0.81	0.76	0.84	
CFMT	McDonald's ω	0.91	0.89	0.92	0.01
	Cronbach's α	0.90	0.89	0.92	0.01
	ICC(2,1)	0.68	0.18	0.85	
CCMT	McDonald's ω	0.84	0.82	0.87	0.01
	Cronbach's α	0.84	0.82	0.87	0.01
	ICC(2,1)	0.66	0.60	0.72	
CFPT	McDonald's ω	0.83	0.80	0.85	0.01
	Cronbach's α	0.82	0.80	0.85	0.01
	ICC(2,1)	0.75	0.70	0.80	

items contribute significantly to the measurement of difficulties associated with developmental prosopagnosia. A detailed summary of the item-level reliability statistics is presented in Table 7.

Table 7. Reliability of the Spanish version of the PI20 if each item was deleted (Study 2).

Item	McDonald's ω	Cronbach's α	Mean	SD
PI20_01	0.829	0.824	2.225	0.932
PI20_02	0.827	0.821	2.291	0.992
PI20_03	0.848	0.846	3.447	0.997
PI20_04	0.828	0.822	2.051	0.973
PI20_05	0.829	0.824	1.586	0.709
PI20_06	0.829	0.824	1.907	0.850
PI20_07	0.821	0.816	1.805	0.932
PI20_08	0.848	0.841	2.685	1.169
PI20_09	0.841	0.835	3.105	0.990
PI20_10	0.831	0.826	2.141	0.872
PI20_11	0.829	0.823	1.790	0.911
PI20_12	0.822	0.817	2.126	0.974
PI20_13	0.839	0.831	1.955	1.018
PI20_14	0.835	0.829	1.877	0.918
PI20_15	0.829	0.823	1.691	0.870
PI20_16	0.828	0.823	1.856	0.946
PI20_17	0.846	0.840	2.796	1.143
PI20_18	0.833	0.828	1.387	0.628
PI20_19	0.850	0.845	2.847	1.043
PI20_20	0.832	0.827	2.111	0.971

Factor Analysis

Confirmatory factor analysis (CFA) was conducted to test the unidimensional structure of the Mexican Spanish version of PI20, as proposed in the original scale. The model, estimated by using diagonally weighted least squares as recommended for ordinal items (Li, 2016), had acceptable fit indices: the Comparative Fit Index (CFI) was .960, and the Tucker–Lewis Index (TLI) was .955, both exceeding the conventional cutoff of .90. The Root Mean Square Error of Approximation (RMSEA) was .089 (90% CI [.081, .096]), and the Standardized Root Mean Square Residual (SRMR) was .078, both near acceptable thresholds (Hu & Bentler, 1999).

The results support the adequacy of a one-factor solution, supporting the interpretation that the items of the PI20 measure a common underlying construct associated with developmental prosopagnosia. Most factor loadings were statistically significant and above the .30 threshold, while a few items (3, 8, 17, and 19) showed lower loadings (< .30). Nonetheless, all items contributed significantly to the model, supporting their retention in the scale. A summary of the model fit indices and parameter estimates is provided in Supplementary Table S3.

An exploratory factor analysis (EFA) was conducted to further examine the underlying structure of the Mexican Spanish version of PI20. The analysis revealed a two-factor solution using a principal axis factoring with oblique rotation. Items such as PI20_07 (*"En ocasiones tengo que advertirles a las personas nuevas que conozco que soy malo reconociendo los rostros*") and PI20_12 ("*Tengo que esforzarme más que otras personas para memorizar rostros*") loaded strongly on factor 1, whereas items such as PI20_08 ("*Se me facilita imaginarme el rostro de alguien en mi mente*") and PI20_13 ("*Confío mucho en mi habilidad para reconocerme en fotografías*") showed higher loadings on factor 2, suggesting the presence of two distinct but strongly related components within the DP construct. Items in factor 2 characterized as items where agreement with the sentence indicated better face recognition abilities. Factors were correlated (r=.330).

These results suggest that the PI20 items capture related aspects of face recognition difficulties that may be organized into partially distinct but correlated components. The factors were moderately correlated (r = .33), indicating that they reflect related processes within the broader construct of developmental prosopagnosia. All factor loadings were statistically significant and above 0.28, indicating that individual items contributed meaningfully to the latent structure. Therefore, the exploratory factor analysis should be interpreted as providing preliminary insights into the internal organization of the scale rather than definitive evidence regarding its dimensionality. A summary of the factor loadings and parameter estimates is presented in Table 8.

Convergent & Discriminant Validity

Evidence for convergent and discriminant validity of PI20 was assessed through correlations with objective performance on face and object recognition tasks. Although these tasks are widely used in research on face processing, they have not been formally validated in Mexican populations. Therefore, the results presented here should be in-

Table 8. Exploratory Factor Analysis of the Spanish version of the PI20 (Study 2).

Item	Factor 1	Factor 2	Uniqueness
PI20_07	0.748		0.425
PI20_12	0.732		0.476
PI20_02	0.648		0.607
PI20_04	0.631		0.626
PI20_11	0.621		0.631
PI20_01	0.606		0.662
PI20_10	0.593		0.678
PI20_16	0.591		0.646
PI20_05	0.589		0.624
PI20_06	0.574		0.643
PI20_15	0.529		0.652
PI20_20	0.517		0.741
PI20_18	0.415		0.74
PI20_14	0.354		0.773
PI20_08		0.478	0.774
PI20_13		0.455	0.722
PI20_09		0.432	0.777
PI20_03		-0.339	0.885
PI20_19		0.305	0.916
PI20_17		0.282	0.898

Note. Uniqueness represents the proportion of variance in each item not explained by the common factor ($1 - h^2$).

terpreted as preliminary evidence based on internationally established measures. As expected, the PI20 showed a significant negative correlation with the Cambridge Face Memory Test (CFMT; ($r(316) = -.229$, $p < .001$)), indicating that individuals who reported greater difficulties in face recognition tended to perform worse on a standardized test of face memory. This correlation aligns with the threshold proposed by Gignac & Szodorai (2016), who consider effect sizes of $r \approx .22$ to reflect moderate associations, thereby supporting the convergent validity of the instrument.

In contrast, the PI20 did not show a significant correlation with errors in the Cambridge Face Perception Test (CFPT; ($r(314) = .047$, $p = .41$)), nor with its accuracy-based scoring ($r(314) = -.139$, $p = .08$) , providing limited support for convergent validity with perceptual tasks. No significant association was observed with the Cambridge Car Memory Test (CCMT; ($r(313) = -.106$, $p = .061$)), providing preliminary support for discriminant validity of the PI20 in relation to non-face object memory. The CCMT was specifically designed to mirror the CFMT in format while assessing memory for non-face objects (cars), thus allowing for a controlled evaluation of discriminant validity (Dennett et al., 2012). These results suggest that the PI20 specifically reflects face recognition difficulties rather than general visual memory performance. All correlations are reported in Table 9.

Conclusion of Study 2

The psychometric evaluation of the adapted Prosopagnosia Index 20-Items (PI20) provides preliminary evidence supporting its reliability and validity for assessing self-reported difficulties in face recognition in the Mexican population. The instrument demonstrated high internal consistency and excellent test-retest reliability. The internal structure was consistent with the original unidimensional model. Convergent validity was supported by a negative correlation with the Cambridge Face Memory Test (CFMT); on the other hand, discriminant validity was established by the absence of correlations with the Cambridge Car Memory Test (CCMT). Items with low factor loadings did not substantially affect internal consistency or reliability, and therefore their inclusion does not represent a limitation for the scale. Overall, the results suggest that the PI20 may be useful for assessing self-reported difficulties in face recognition in Mexican Spanish-speaking contexts. However, because the present study did not include participants with independently confirmed diagnoses of developmental prosopagnosia, the findings should not be interpreted as evidence of diagnostic validity.

General Discussion

This research adapted and provided preliminary validation evidence for the PI20 for use in Mexican Spanish for the screening of DP, following a rigorous, multi-phase process that ensured semantic, conceptual, and cultural equivalence. In Study 1, cognitive interviews and pilot testing with native Spanish speakers from Mexico enabled the refinement of specific items, particularly items 8, 18, and 20, to enhance clarity and contextual relevance. The results supported the content validity of the adaptation and laid a solid foundation for subsequent psychometric evaluation. In Study 2, the adapted instrument showed preliminary evidence of adequate psychometric properties in a larger sample. The PI20 showed high internal consistency, excellent test-retest reliability, and a factorial structure consistent with the original unidimensional model. Evidence of convergent validity was established through a negative correlation with the Cambridge Face Memory Test (CFMT), while discriminant validity was confirmed by the absence of significant correlations with the Cambridge Car Memory Test (CCMT), an object recognition task. Exploratory factor analysis suggested the presence of two correlated components, though the overall structure remained consistent with a unidimensional interpretation, as confirmed by confirmatory factor analysis.

These findings provide preliminary evidence supporting the reliability and validity of the Mexican Spanish version of the PI20 for identifying difficulties associated with DP among Mexican Spanish-speaking populations in Mexico. In this context, the PI20 should be interpreted primarily as a measure of individual differences in self-reported face recognition difficulties within the general population, rather than as a standalone diagnostic instrument for developmental prosopagnosia. Its psychometric properties, combined with evidence supporting its linguistic and con-

Table 9. Correlations between the Spanish version of the PI20 with CFMT, CFPT and CCMT (CFMT, face memory; CCMT, car memory; CFPT, face perception; Study 2).

	PI20_Total	CFMT_Total	CFPT_total_errors
CFMT_Total	-0.229 (<.001)		
CFPT_total_errors	0.047 (.407)	-0.481 (<.001)	
CCMT_Total	-0.106 (.061)	0.283 (<.001)	-0.257 (<.001)

Note. Each cell reports: Pearson's r (p-value).

textual appropriateness, suggest that it may be useful for research and screening purposes. The adaptation process followed internationally recognized guidelines, including those from the (International Test Commission, 2017), and involved forward independent translations, expert committee review, and iterative refinement based on cognitive interviews and pilot testing. As in other adaptations, such as the Danish and French versions, cultural and linguistic challenges emerged—particularly in items requiring contextual reinterpretation (e.g., item 8)—and were resolved through methodologically rigorous procedures (Sexton et al., 2023). The development of culturally adapted diagnostic instruments is essential for advancing the assessment of neurocognitive conditions in diverse populations.

The use of a mixed-methods, cross-sectional design enabled a comprehensive evaluation of the instrument. In Study 1, qualitative insights from cognitive interviews provided content validity evidence, ensuring the clarity and cultural relevance of the translated items. In study 2, psychometric analyses based on large-scale administration provided preliminary support for the instrument's structural, convergent, and discriminant validity. This methodological integration aligns with best practices in test adaptation and validation (Arizpe et al., 2019; Hambleton & Lee, 2013), resulting in a version of the PI20 that is both psychometrically sound and contextually appropriate.

The average scores of the Spanish version of the PI20 were highly similar to those reported by Shah et al. (2015). Particularly, the mean PI20 score obtained in this study was comparable to their control sample (42 vs 38.90), and the pattern of the means for each item was also equivalent. For example, the items 3, 9, and 19 (3.69, 2.76, and 2.79, respectively), had the highest means in Shah et al. (2015), and also were the highest in this study (3.45, 3.11, and 2.85, respectively). Therefore, the descriptive results obtained with the Spanish version of the PI20 in the Mexican sample are also evidence of the equivalence to the original English version.

Reliability analyses revealed excellent internal consistency, with McDonald's Omega and Cronbach's Alpha both at .84—comparable to those of the original scale (Shah et al., 2015) and other adaptations (Nørkær et al., 2023). Item-level analysis showed no substantial increase in reliability upon the removal of any item, supporting the internal coherence of the scale. Test-retest reliability, assessed over a six-week interval, produced an intraclass correlation coefficient (ICC) of .81, consistent with standards for temporal stability (Koo & Li, 2016).

Construct validity was supported by both confirmatory and exploratory factor analyses, which indicated a predominantly unidimensional structure. This aligns with the theoretical basis of the PI20 and previous validation studies. Evidence for convergent validity was provided by a moderate negative correlation between PI20 scores and CFMT performance ($r = -.229$), matching the threshold for moderate effects proposed by Gignac and Szodorai (2016) and consistent with prior findings (Gray et al., 2017). Discriminant validity was demonstrated by the absence of a significant correlation with the Cambridge Car Memory Test (CCMT), confirming that the PI20 specifically measures face recognition difficulties rather than general object memory.

The weak and nonsignificant correlations observed between PI20 and the Cambridge Face Perception Test (CFPT) suggest that face perception and face memory may represent partially dissociable cognitive processes. This pattern is consistent with previous findings reported by DeGutis et al. (2023) and Nørkær et al. (2023), and supports the relevance of considering subtypes of developmental prosopagnosia—such as apperceptive and associative forms—in both research and clinical assessments.

The current adaptation of the PI20 contributes to this goal and paves the way for future studies exploring its applicability across broader demographic groups, including children, older adults, and clinical populations. Additionally, examining sociocultural variables that may influence face perception can further inform the contextualization of prosopagnosia and guide tailored intervention strategies. For example, prior studies, such as Nigrou et al. (2024), have found associations between higher PI20 scores and increased symptoms of social anxiety, implying that face recognition deficits may have broader psychosocial implications. Although such variables were not examined in the present study, they represent potential directions for future research within the Mexican population.

Additionally, although the CFMT, CFPT, and CCMT are widely used internationally as measures of face perception and memory, these tasks have not yet been formally validated in Mexican populations. It is important to note that the present validation was conducted with Mexican participants recruited primarily in the city of Tijuana, a border region characterized by high levels of internal migration and frequent exposure to bilingual and bicultural environments. These sociocultural characteristics may differ from those present in other regions of Mexico. Therefore, caution is warranted when generalizing the present findings to other Mexican regions or to Spanish-speaking populations with different linguistic and cultural contexts, and further studies will be necessary to evaluate the applicability of the scale in those populations.

In conclusion, the Mexican Spanish version of the 20-Item Prosopagnosia Index (PI20) provides promising preliminary evidence of reliability and validity as a self-report measure of face recognition difficulties in Mexican Spanish-speaking populations. Although additional validation studies are needed, particularly involving objective face-processing measures formally validated in Mexican populations, the present study represents an important first step toward the development of psychometrically sound assessment tools for developmental prosopagnosia research in Mexico. Future studies may extend this work by exploring its applicability across broader demographic groups and by examining the psychosocial and cognitive correlates of face recognition difficulties in greater depth.

Limitations and Future Directions

Several limitations of the present study should be acknowledged. First, although the total sample size used for the psychometric analyses was adequate for correlational procedures and confirmatory factor analysis, the subgroup used for the test-retest analysis was relatively small ($n = 31$). Consequently, estimates of test-retest reliability should be interpreted as preliminary evidence rather than conclusive evidence of long-term reliability. Future studies with larger longitudinal samples will be necessary to obtain more stable estimates of test-retest reliability in diverse Mexican populations.

Second, convergent and discriminant validity were assessed using adaptations of the Cambridge Face Memory Test (CFMT), the Cambridge Face Perception Test (CFPT), and the Cambridge Car Memory Test (CCMT). Although these instruments are widely used internationally in research on facial processing, they have not yet undergone formal psychometric validation procedures in a Mexican population. Therefore, the correlations reported in the present study should be interpreted with caution as preliminary evidence based on adaptations of internationally established measures, rather than as conclusive evidence of validity within the Mexican cultural context. Future research should prioritize the validation of objective measures of facial processing for Spanish-speaking Mexican populations, in order to strengthen the assessment of facial recognition difficulties and related constructs.

It is important to note that this study should be viewed as an initial, albeit necessary, step toward the development of psychometrically sound instruments for assessing facial recognition difficulties in the Mexican population. Despite these limitations, the study provides preliminary evidence supporting the reliability and validity of the Spanish-language Mexican version of the PI20 and establishes a methodological foundation for future validation research in this area.

Author Contributions

Contributed to conception and design: MAM, AC, VL, ABL

Contributed to acquisition of data: AC, MAM, VL

Contributed to analysis and interpretation of data: MAM, AC, VL, ABL

Contributed to supervision, and data curation: MAM

Drafted and/or revised the article: MAM, AC, VL

Approved the submitted version for publication: MAM, AC, VL

Acknowledgements

Data collection: Alam Ibarra, Javier Samaniego Ojeda, Emily Sayuri Morales Castelan, Azul Gabriela Camacho Hernandez, Maria de la Luz Perfecto Sanchez, & Ximena Estephanya Velasco Lopez. **Drafts revisions**: Alam Ibarra. **Translators**: Ricardo Eliud Ramirez, Genesis Cerrillo, & Esmeralda Mosqueda. **Expert panel**: Hernán Partida Ramírez, Víctor Fuentes-Barradas & Diana Estefanía Andrade.

Funding Information

MM received funding from the Institutional Research Coordination (Project #517) of CETYS University.

Ethics Statement

Ethical approval was obtained from the Research Ethics Committee of CETYS University (#D-CEI154).

Competing Interests

Authors declare no competing interests for this study.

Data Accessibility Statement

All the translation materials, anonymized participant data, and analysis scripts can be found on this paper's project repository: https://osf.io/k8vwt/

Editors: Daniel P. Moriarity (Associate Editor)

Submitted: September 24, 2025 PDT. Accepted: May 18, 2026 PDT. Published: June 24, 2026 PDT.

References

Anzellotti, S., & Caramazza, A. (2014). The neural mechanisms for the recognition of face identity in humans. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00672

Arifin, W. nor. (2026). *Sample Size Calculator* [Web]. http://wnarifin.github.io

Arizpe, J. M., Saad, E., Douglas, A. O., Germine, L., Wilmer, J. B., & DeGutis, J. M. (2019). Self-reported face recognition is highly valid, but alone is not highly discriminative of prosopagnosia-level performance on objective assessments. *Behavior Research Methods*, *51*(3), 1102–1116. https://doi.org/10.3758/s13428-018-01195-w

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Benton, A., Sivan, A. B., Hamsher, K. deS., Varney, N. R., & Spreen, O. (1994). *Contributions to Neuropsychological Assessment: A clinical manual* (2nd ed.). Oxford University Press.

Brewer, R., Gray, K., & Cook, R. (2018). Should developmental prosopagnosia, developmental body agnosia, and developmental object agnosia be considered independent neurodevelopmental conditions? *Cognitive Neuropsychology*, *35*, 59–62. https://doi.org/10.1080/02643294.2018.1433153

Champely, S. (2020). *pwr: Basic Functions for Power Analysis* (1.3-0) [Dataset]. https://doi.org/10.32614/CRAN.package.pwr

Corrow, S. L., Dalrymple, K. A., & Barton, J. J. S. (2016). Prosopagnosia: Current perspectives. *Eye and Brain*, *8*, 165–175. https://doi.org/10.2147/EB.S92838

DeGutis, J., Bahierathan, K., Barahona, K., Lee, E., Evans, T. C., Shin, H. M., Mishra, M., Likitlersuang, J., & Wilmer, J. B. (2023). What is the prevalence of developmental prosopagnosia? An empirical assessment of different diagnostic cutoffs. *Cortex*, *161*, 51–64. https://doi.org/10.1016/j.cortex.2022.12.014

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2012). The Cambridge Car Memory Test: A task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. *Behavior Research Methods*, *44*(2), 587–605. https://doi.org/10.3758/s13428-011-0160-2

Duchaine, B., Germine, L., & Nakayama, K. (2007). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology*, *24*(4), 419–430. https://doi.org/10.1080/02643290701380491

Duchaine, B., & Nakayama, K. (2006). Developmental prosopagnosia: A window to content-specific face processing. *Current Opinion in Neurobiology*, *16*(2), 166–173. https://doi.org/10.1016/j.conb.2006.03.003

Garrido, L., Duchaine, B., & Nakayama, K. (2008). Face detection in normal and prosopagnosic individuals. *Journal of Neuropsychology*, *2*(1), 119–140. https://doi.org/10.1348/174866407X246843

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Gray, K. L. H., Bird, G., & Cook, R. (2017). Robust associations between the 20-item prosopagnosia index and the Cambridge Face Memory Test in the general population. *Royal Society Open Science*, *4*(3), 160923. https://doi.org/10.1098/rsos.160923

Hambleton, R. K., & Lee, M. K. (2013). Methods for Translating and Adapting Tests to Increase Cross-Language Validity. In D. H. Saklofske, C. R. Reynolds, & V. Schwean (Eds.), *The Oxford Handbook of Child Psychological Assessment* (Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199796304.013.0008

Hambleton, R. K., & Zenisky, A. L. (2010). Translating and Adapting Tests for Cross-Cultural Assessments. In D. Matsumoto & F. J. R. Van De Vijver (Eds.), *Cross-Cultural Research Methods in Psychology* (1st ed., pp. 46–70). Cambridge University Press. https://doi.org/10.1017/CBO9780511779381.004

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Instituto Nacional Electoral (INE). (2020). *Anexo K: Análisis sociodemográfico de Baja California 2020. [Sociodemographic analysis of Baja California 2020].* https://repositoriodocumental.ine.mx/xmlui/bitstream/handle/123456789/177892/Anexo-K-Analisis-sociodemografico-de-Baja-California-2020.pdf

International Test Commission. (2017). *ITC Guidelines for Translating and Adapting Tests* (Version 2.4, 2nd ed.). International Test Commission. http://www.InTestCom.org

Kim, K. H. (2005). The Relation Among Fit Indexes, Power, and Sample Size in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 368–390. https://doi.org/10.1207/s15328007sem1203_2

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Liljequist, D., Elfving, B., & Skavberg Roaldsen, K. (2019). Intraclass correlation – A discussion and demonstration of basic features. *PLOS ONE*, *14*(7), e0219854. https://doi.org/10.1371/journal.pone.0219854

Maw, K. J., Burns, E. J., & Beattie, G. (2025). Prosopagnosia is highly comorbid in individuals with probable developmental coordination disorder (DCD). *Quarterly Journal of Experimental Psychology*, *78*(8), 1501–1522. https://doi.org/10.1177/17470218241275977

Nakashima, S. F., Ukezono, M., Sudo, R., Nunoi, M., Kitagami, S., Okubo, M., Toriyama, R., Morimoto, Y., & Takano, Y. (2020). Development of a Japanese version of the 20-item prosopagnosia index (PI20-J) and examination of its reliability and validity. *The Japanese Journal of Psychology*, *90*(6), 603–613. https://doi.org/10.4992/jjpsy.90.18235

Nguyen, C. M., Rampa, S., Staios, M., Nielsen, T. R., Zapparoli, B., Zhou, X. E., Mbakile-Mahlanza, L., Colon, J., Hammond, A., Hendriks, M., Kgolo, T., Serrano, Y., Marquine, M. J., Dutt, A., Evans, J., & Judd, T. (2024). Neuropsychological application of the International Test Commission Guidelines for Translation and Adapting of Tests. *Journal of the International Neuropsychological Society*, *30*(7), 621–634. https://doi.org/10.1017/S1355617724000286

Nigrou, T., Hansenne, M., & Devue, C. (2024). Exploration of the Links Between Psychosocial Well-being and Face Recognition Skills in a French-Speaking Sample. *Psychologica Belgica*, *64*(1). https://doi.org/10.5334/pb.1294

Nørkær, E., Guðbjörnsdóttir, E., Roest, S. B., Shah, P., Gerlach, C., & Starrfelt, R. (2023). The Danish Version of the 20-Item Prosopagnosia Index (PI20): Translation, Validation and a Link to Face Perception. *Brain Sciences*, *13*(2), 337. https://doi.org/10.3390/brainsci13020337

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Posit team. (2023). *RStudio: Integrated Development Environment for R*. Posit. Software, PBC. http://www.posit.co/

Reynolds, C. R., & Suzuki, L. A. (2012). Bias in Psychological Assessment: An Empirical Review and Recommendations. In I. Weiner (Ed.), *Handbook of Psychology, Second Edition* (1st ed.). Wiley. https://doi.org/10.1002/9781118133880.hop210004

Sexton, O., Pilley, S., d'Ardenne, J., & Bull, R. (2023). *Cognitive Interviewing and what it can be used for*. https://www.ncrm.ac.uk/resources/online/all/?id=20816

Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): A self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science*, *2*(6), 140343. https://doi.org/10.1098/rsos.140343

Shoukri, M. M., Asyali, M. H., & Donner, A. (2004). Sample size requirements for the design of reliability study: Review and new results. *Statistical Methods in Medical Research*, *13*(4), 251–271. https://doi.org/10.1191/0962280204sm365ra

Smith, C., & Susilo, T. (2021). Normal colour perception in developmental prosopagnosia. *Scientific Reports*, *11*(1), 13741–13741. https://doi.org/10.1038/s41598-021-92840-6

Stensen, K., & Lydersen, S. (2022). Internal consistency: From alpha to omega? *Tidsskrift for Den Norske Legeforening*. https://doi.org/10.4045/tidsskr.22.0112

Stumps, A., Saad, E., Rothlein, D., Verfaellie, M., & DeGutis, J. (2020). Characterizing developmental prosopagnosia beyond face perception: Impaired recollection but intact familiarity recognition. *Cortex*, *130*, 64–77. https://doi.org/10.1016/j.cortex.2020.04.016

Sun, W., Wang, Y., Wang, J., & Luo, F. (2021). Psychometric Properties of the Chinese version of the 20-item Prosopagnosia Index (PI20). *E3S Web of Conferences*, *271*, 01036. https://doi.org/10.1051/e3sconf/202127101036

Tagliente, S., Passarelli, M., D'Elia, V., Palmisano, A., Dunn, J. D., Masini, M., Lanciano, T., Curci, A., & Rivolta, D. (2023). Self-reported face recognition abilities moderately predict face-learning skills: Evidence from Italian samples. *Heliyon*, *9*(3), e14125. https://doi.org/10.1016/j.heliyon.2023.e14125

Towler, J., Fisher, K., & Eimer, M. (2017). The cognitive and neural basis of developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, *70*(2), 316–344. https://doi.org/10.1080/17470218.2016.1165263

Ventura, P., Livingston, L. A., & Shah, P. (2018). Adults have moderate-to-good insight into their face recognition ability: Further validation of the 20-item Prosopagnosia Index in a Portuguese sample. *Quarterly Journal of Experimental Psychology*, *71*(12), 2677–2679. https://doi.org/10.1177/1747021818765652