



OPEN

Clustering of neuropsychological traits of preschoolers

Mario Treviño^{2✉}, Beatriz Beltrán-Navarro³, Ricardo Medina-Coss y León² & Esmeralda Matute^{1✉}

Neuropsychological tests (targeting cognitive, linguistic, motor, and executive abilities) are grouped in neuropsychological domains that are thought to be stable through adulthood. However, this assumption does not always hold true, particularly during young children's early developmental phase. Here, we explored how the neuropsychological profile of typical Spanish-speaking preschoolers varied and consolidated with age. We recruited 643 monolingual Latin-American children from Mexico, Colombia, and Guatemala, with ages spanning from 30 to 71 months of age, and applied a novel neuropsychological examination which combined a total of 52 tests covering five classical neuropsychological domains: receptive, expressive, attention/memory, processing, and executive functions. These tests' scores uncovered a correlational structure across neuropsychological functions that could not be explained by chance. Notably, these correlations' overall strength, but not their interdependence across domains, dramatically increased with age. Moreover, by applying conventional clustering techniques to classify the experimental data, we found a stable representation of two clusters of children with distinctive traits, with cultural factors contributing to this classification scheme. We also found that the tasks were well organized in a network of abilities, where nodes with highest highest interconnectedness were those that required multimodal processing. These results contribute to our understanding of children's 'normal' development and could help identify how failure in particular functions forecasts the emergence of neurodevelopmental disorders. Our analytic methods might become useful to characterize individual differences and improve educational practices and interventions.

Historically, neuropsychological research has focused on the administration of tests to identify brain dysfunctions and learning disabilities. The traditional view, mainly derived from lesion studies, proposes that the human brain has discrete and separate modules specializing in specific functions¹. Under this notion, it should be possible to associate a localized brain lesion with a single functional disruption. A more contemporary perspective, however, proposes that the brain can not be assumed to be modular because it is not composed of isolated parts. Indeed, evidence shows that the brain exhibits large-scale interconnectivity patterns that are incompatible with the idea of independently functioning modules. Accordingly, modern neuroscience tries to explain behavior due to complex interactions across multiple brain areas². For example, executive functions, which depend on the prefrontal cortex and subcortical systems, are generally described as multidimensional constructs consisting of several interrelated sub-functions³⁻⁵. These sub-functions include volition, goal selection, planning, inhibition, cognitive flexibility, working memory, and evaluating the results of actions, among others. Neuroconstructivist and developmental approaches also view the brain as a dynamic self-organized system with internal (physiological, genetic) and external (context) factors that shape its maturation, also with patterns of activation becoming specialized through this process¹. Therefore, almost every aspect of development turns out to be dynamic and interactive⁶.

Many studies have shown how neuropsychological functions unfold from infancy to adolescence, within different hierarchical modules^{3,7-9}. These modules develop at different rates, starting on subcortical areas that allow arousal, followed by the maturation of regions participating in motor and sensory functions. After that, circuits involved in motor coordination, spatial orientation, and language develop. Similarly, the development of integrative and multimodal functions associated with tertiary sensorial areas gives rise to abilities such as reading, writing, and arithmetic skills. The prefrontal cortex is the last area to mature: it participates in executive functions

¹Laboratorio de Neuropsicología y Neurolingüística, Instituto de Neurociencias, Universidad de Guadalajara, Guadalajara, México. ²Laboratorio de Plasticidad Cortical y Aprendizaje Perceptual, Instituto de Neurociencias, Universidad de Guadalajara, Guadalajara, Jalisco, México. ³Departamento de Neurociencias, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, Guadalajara, Jalisco, México. ✉email: mariomtv@hotmail.com; ematute@encar.udg.mx

necessary for goal-directed behavior¹⁰. Consequently, some impairments in any of these functions/regions are associated with a range of disorders and severities according to the neurodevelopmental phase when they occur.

Trying to understand the brain as a complex system gives rise to novel questions. How do cognitive functions change and interact over developmental time, and what are the internal and external elements that affect such cognitive development? Consider, for instance, the development of language. Here, the maturation of brain areas involved in motor control (secondary motor areas) allows the production of phonemes. Isolated words and sentences are then produced in coordination with secondary motor areas. These abilities are the precursors of storytelling, which requires a complex interaction between memory capacity and the presence of semantic and morphosyntactic components¹¹. Reading and writing appear around 6 years of age as multimodal functions, where complex metalinguistic processes arise. Finally, the maturation of the prefrontal lobes unfolds as a multi-stage process, with distinctive abilities improving in different ways and at different rates until reaching maturity in adulthood¹². This maturation process allows planning skills, working memory, self-reflexive processes (crucial for written expression), argumentation, and critical thinking.

All in all, although the brain completes its development until adulthood, the neurodevelopmental processes during childhood are markedly different across ages. For example, there is a high processing variability across multiple neuropsychological domains in children below 6 years of age, particularly in executive functions. These early developmental phases are highly relevant because they impact numerous functional areas, such as school and social performance. Furthermore, the immaturity of neuropsychological functions in young children implies that they can be vulnerable to cerebral insults. For all these reasons, it is crucial to characterize whether and how relevant neuropsychological functions change during the early stages of child development.

Here, we explored how the neuropsychological profile of typical monolingual Spanish-speaking Latin-American preschoolers varied and consolidated with age. Our main aims were to: (1) track the neuropsychological development in these preschoolers, (2) determine how domestic and educational factors were related to their neuropsychological profiles, and (3) explore whether and how these abilities were organized and interrelated. We employed a battery of 52 tests, grouped into five neuropsychological functions: receptive, expressive, attention/memory, processing, and executive ones. This battery was developed following the conceptual framework of Lezak et al.⁵, and standardized for Spanish-speaking Latin-American children¹³. Experienced psychologists carried out testing, and the children were organized into seven age groups, ranging from 30 to 71 months of age. Our analytic results revealed well-orchestrated changes across age groups that were influenced by educational and other socio-cultural variables. Interestingly, even though our participants were typical middle-class preschoolers, they could be classified into two well-differentiated clusters. Our analysis also revealed that the neuropsychological abilities conformed a network of well-interconnected nodes. Within this network, the tests belonging to expressive, processing, and executive functions exhibited the highest interconnectedness. Our results and methods serve to understand the development of neuropsychological functions and how individual differences emerge; they may be useful for educational practices and interventions.

Results

Classification of Latin-American children across neuropsychological functions. Typically, examiners use neuropsychological tests (covering cognitive, linguistic, motor, and executive domains) to explore discrete responses associated with particular functions. First, we wondered if such apparently discrete abilities (responses) do not occur in isolation but, instead, they contribute to more complex skills. For this, we characterized the distribution of neurocognitive abilities in Latin-American children from 30 to 71 months of age. We applied a battery of 52 tests (Supplementary Table 1) covering receptive (8 tests), expressive (15), attention/memory (9), processing (13), and executive functions (7), on a total of 643 typical children (Supplementary Table 2). We merged the data from female and male children because there were no gender differences when comparing their averaged scores (two-tailed Wilcoxon rank-sum test; $P=0.15$). These tests' normalized scores revealed a strong task-dependency, yet with a high heterogeneity across children (Fig. 1a). We thus wondered how the scores obtained from the different tests were related to domains within the same and other neuropsychological functions. Intuitively, one would expect similar scores among tests belonging to the same neuropsychological function. We performed a correlation analysis across all tasks using the entire database and tested the observed correlations' significance. We did this by creating surrogate comparison data ($n=1000$) by shuffling each participant's scores across the 52 tests. In Fig. 1b, we illustrate the cross-correlation matrix derived from this analysis with significant relationships marked as colored squares and non-significant (n.s.) values appearing as empty squares (i.e., $P\geq 0.05$). This matrix reveals a quite heterogeneous structure with strong and weak relationships among tests that belong to practically all functions (represented with different color bars to the left and above the correlation matrix). Interestingly, tests from attention/memory functions (in orange) exhibited smaller correlations than the rest of the tests (one-tailed Wilcoxon rank-sum test; $P\leq 0.027$ for all cases). To explore alternative relational schemes among tests, we re-organized the correlation matrix by sorting its columns by the summed correlations across all tests (from highest to lowest; Fig. 1c). The resulting heterogeneous distribution of functions (to which the sorted tests belonged) reveals intermingled functions for strongly correlated tests (illustrated in the color bar on top of the correlation matrix).

Next, we wondered whether the high heterogeneity in the children's scores across tasks could reflect particular sub-groups of children with specific skills within the sampled population. We applied a nearest neighbor classification technique using the Euclidean distance as a measure for cluster center initialization. Using the Calinski-Harabasz and the Davies-Bouldin criteria (Supplementary Fig. 1a–e), we found that the original data set could be organized into two main clusters (C_1 and C_2 ; Fig. 1d). Using the same optimization criteria, we found that these two clusters could be further subdivided into two additional sub-clusters each (Fig. 1e and Supplementary Fig. 1f,g). The two main clusters differed in that cluster 1 (C_1) was associated with ~47% lower average scores

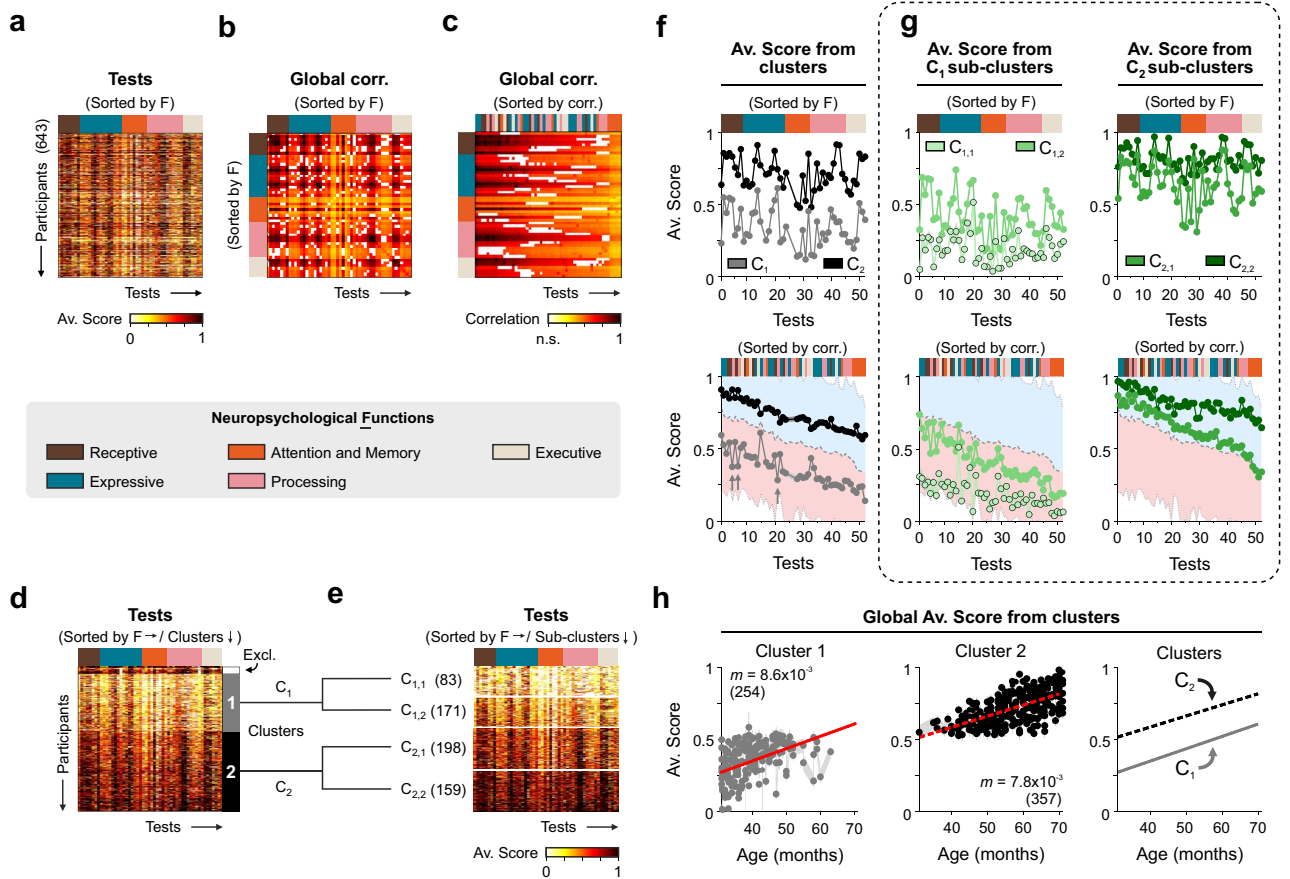


Figure 1. Classifying Latin-American children’s performance across distinct tasks and neuropsychological functions. (a) Scores from a battery of 52 neuropsychological tests performed on 643 Latin-American children. Grouping of tests into receptive (walnut brown), expressive (mosaic blue), attention/memory (orange), processing (flamingo pink), and executive (cloud cream yellow) functions (color bar on the bottom of the heatmap). (b) Peak global cross-correlations across tasks sorted by functions or (c) by the summed correlations across tasks for the entire data set. Filled pixels represent values that cannot be explained by chance ($P < 0.05$). (d) Classification of children into two main clusters (C_1 and C_2). (e) Further partition of the two main clusters into two sub-clusters each ($C_{1,1}$, $C_{1,2}$, $C_{2,1}$, and $C_{2,2}$). (f) Average scores \pm S.E.M. for the groups of children from the main clusters sorted by neuropsychological functions (top) or by the summed correlations (bottom). Gray arrows point to three tests that exhibited average values below those expected if they followed the trend of scores sorted by correlation (test #5: Left hand recognition; #7: Following oral instructions; #21: Naming body parts). (g) Average scores from each sub-cluster sorted by neuropsychological functions and by the summed correlations (same arrangement as in f). Shaded areas in panels F and G depict the regions covering 1.5 standard deviations away from the mean of the population, which is shown as a grey dotted line. (h) Global scores (i.e., the average of all tests) as a function of age for C_1 and C_2 . Linear fits display similar slopes but different intercepts for each cluster. Number of participants in parentheses.

across all tasks compared to cluster 2 (C_2 ; scores C_1 : $34.15 \pm 0.84\%$, $n = 254$; scores C_2 : $71.44 \pm 0.57\%$, $n = 357$; Kruskal–Wallis multi-comparison test with Bonferroni *posthoc* correction; $P < 0.0001$ for all tests; Fig. 1f). Practically all functions were similarly reduced in C_1 relative to C_2 (i.e., they were scaled-down). Only three tests exhibited average values below those expected if they followed the trend of scores sorted by correlation (Tests #5: Left-hand recognition; #7: Following oral instructions; #21: Naming body parts; depicted with gray arrows in Fig. 1f). We also found a similar pattern of distinctive average scores across all tests for the corresponding sub-clusters ($C_{1,1}$: $18.65 \pm 1.10\%$, $n = 83$; $C_{1,2}$: $41.68 \pm 0.51\%$, $n = 171$; $C_{2,1}$: $63.36 \pm 0.38\%$, $n = 198$; $C_{2,2}$: $81.49 \pm 0.49\%$, $n = 159$; KW test with Bonferroni *posthoc* correction; $P < 0.0001$; Fig. 1g). Suggesting a graded transition, the

correlation in average scores across all these four classifications revealed stronger similarities between contiguous clusters (not illustrated).

Finally, we searched for potential differences across ages in the neuropsychological functions between the two main clusters. We plotted the average scores from all tests against the age of the children from each group. Interestingly, these two clusters had different intercepts (b) but similar slopes (m ; C_1 : $m = (8.6 \pm 0.5) \times 10^{-3}$, $b = (1.2 \pm 0.5) \times 10^{-3}$; $P < 0.001$; C_2 : $m = (7.8 \pm 0.4) \times 10^{-3}$, $b = (274.6 \pm 2.4) \times 10^{-3}$; Kruskal–Wallis test with Bonferroni *posthoc* correction; $P < 0.001$; Fig. 1h). This result suggests that both clusters exhibited similar improvements with age (as inferred by the slopes of these regressions). Still, each of them was associated with a different performance level. The different intercepts could be explained by external (*v.gr.*, environmental) or internal (*v.gr.*, genetic) conditions related to each group.

Our overall sample of Latin-American children was composed of 76.92% Mexican ($n = 470$), 14.24% Colombian ($n = 87$), and 8.84% Guatemalan ($n = 54$) children (Supplementary Fig. 1h). The difference in the number of participants from each country creates the potential concern that the optimal number of clusters that we found for the entire dataset could be sensitive to educational or cultural differences among these nationalities. We adjusted and compared the linear regressions to the children's global performance by their nationalities (i.e., the average score from all the tests/participants) as a function of age (Supplementary Fig. 1i). We made size-matched linear regressions from sub-sampled groups of children (20–90% with 10% increments, 1000 iterations/category) and found relevant differences in the slopes and intercepts among groups (Supplementary Fig. 1j). More specifically, Mexican children displayed slightly faster increases in their neuropsychological functions (i.e., higher slope) than Colombian and Guatemalan children, yet, the Colombian children had a higher intercept than the other two groups (Mexico: $m = (1.62 \pm 0.05) \times 10^{-2}$, $b = -(22.42 \pm 2.77) \times 10^{-2}$; Colombia: $m = (1.43 \pm 0.14) \times 10^{-2}$, $b = -(7.88 \pm 6.54) \times 10^{-2}$; Guatemala: $m = (1.38 \pm 0.20) \times 10^{-2}$, $b = -(20.67 \pm 11.79) \times 10^{-2}$; Kruskal–Wallis multi-comparison test with Bonferroni *posthoc* correction; $P < 0.0001$, for all cases; Supplementary Fig. 1k). We confirmed these differences by re-calculating the slopes and intercepts from size-matched groups (Supplementary Fig. 1l). Given these results, we then tested whether nationality differences could compromise the entire dataset's sub-clustering structure. We repeated the clustering analysis on the Mexican children only and found a similar classification structure as the one observed for the whole population (of Latin-American children). Namely, on an initial partition, we confirmed the existence of two main clusters, which could be further divided into two sub-clusters each (using Calinski–Harabasz and the Davies–Bouldin criteria; Supplementary Fig. 1m,n). These results confirm the notion that the scores obtained from Latin-American children solving our 52 tests can be organized into four sub-clusters, which exhibit a graded transition in performance across tests.

Development of neuropsychological functions. Neuropsychological functions develop from infancy into adulthood^{14,15}. We wondered if the children's scores for all tests improved similarly with age since we considered high performance as an indirect developmental measure. We used a colormap to illustrate the age-related increases across neuropsychological functions and found distinctive neurodevelopmental profiles (colormaps in Fig. 2a). Since the group averaged scores from our tests presented such a strong dependency on age (lower panel in Fig. 2a), we wondered whether the correlation structure across all tasks could involve a strong neurodevelopmental control. For this, we computed the correlation matrices across tasks from children belonging to each of the seven age groups (ranges: 30–35, 36–41, 42–47, 48–53, 54–59, 60–65, and 66–71 months), and confirmed that the overall correlations across tests increased and became more selective with age (Repeated Measures One-Way ANOVA test, $F = 29.32$, $P = 0.001$; Fig. 2b, Supplementary Fig. 2a). To get an overall estimation of how these correlations varied with age, we calculated the average correlation across all tests as a function of the age group for members from clusters C_1 (middle panel of Fig. 2b) and C_2 (lower panel of Fig. 2b). Tests correlations increased with age for both groups (RM One-Way ANOVA test, C_1 : $F = 52,207.75$, $P < 0.001$; C_2 : $F = 33,168.66$, $P < 0.001$), yet children belonging to C_2 had, on average, a ~ 2.5 higher performance than children from C_1 (middle and lower panels in Fig. 2b; C_1 : $m = 1.6 \pm 0.1$, $b = 5.04 \pm 0.01$; C_2 : $m = 4.2 \pm 0.1$, $b = -1.92 \pm 0.1$; KW-test, $P < 0.005$ for slopes and intercepts).

To estimate if the within-group differences decreased with age (*v.gr.* as one would expect for correlations reaching an asymptote), we calculated the first derivative of the average correlation by taking the difference between average correlations across matrices from contiguous age groups. We found no clear reduction in these numbers through age groups (RM One-Way ANOVA test, $F = 0.47$, $P \geq 0.05$, for both cases; Fig. 2c and Supplementary Fig. 2d), suggesting an un-finished developmental process. These analyses reveal that children from C_2 exhibited a higher increase in their average correlation profile than children from C_1 . Thus, in principle, children from C_2 could be maturing their cognitive functions earlier than children from C_1 . Finally, we also explored if the cluster centroids (and/or the difference between them) varied between age groups. We used a colormap to illustrate the values associated with the cluster centroids. We found, as expected, higher average scores in older age groups (Supplementary Fig. 3a,b). However, the absolute difference between cluster centroids revealed widespread effects across neuropsychological functions with no evident age-related change (Supplementary Fig. 3c). This result is consistent with the fact that both groups exhibited similar changes in their neuropsychological functions through age (Fig. 1h).

Domestic and educational factors linked to neuropsychological classification. Biological (*v.gr.* weeks of gestation, birth weight, head circumference^{16,17}, and environmental (*v.gr.* Family background, marital histories, socio-economical levels^{12,18–20}) factors influence children's neuropsychological development. Since we controlled prenatal factors in our sample (*v.gr.* weeks of gestation; see “Methods”), we next explored whether and how some domestic (household as well as parents age, and educational level) and educational (type of school) factors were related to the observed neuropsychological functions and corresponding classification schemes of

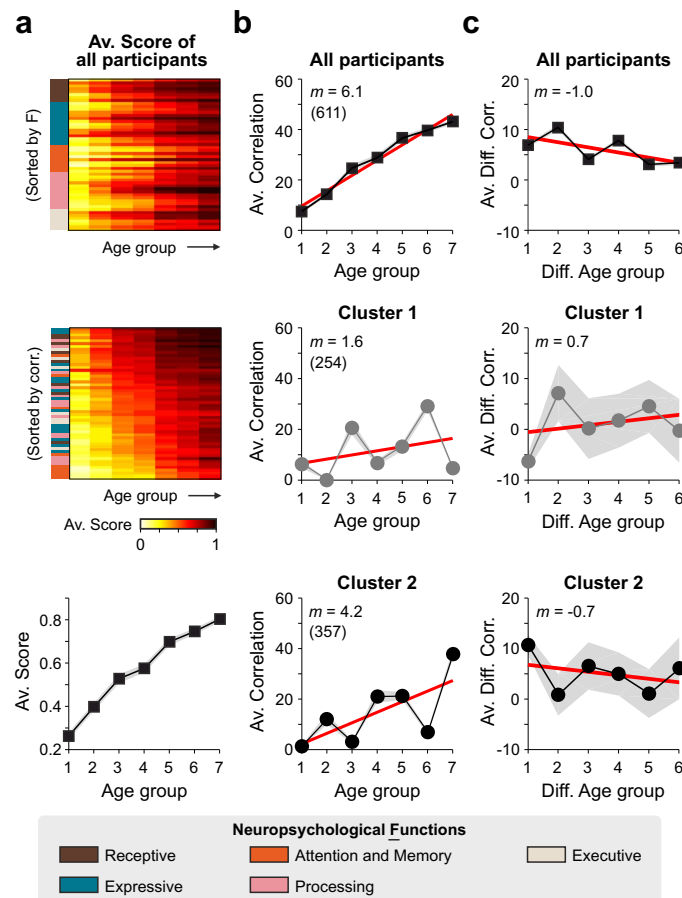


Figure 2. Differential developmental control of neuropsychological functions between identified clusters. (a) Average scores as a function of age group sorted by functions (top) or by increasing correlations (middle). The lower panel shows the average scores from all 52 tests. Averaged (b) and averaged differences (among contiguous age groups; c) in cross-correlations for all tests as a function of the age group for all participants (top), and participants from cluster 1 (center), and cluster 2 (bottom). Red lines depict linear fits with corresponding slopes displayed on the top of each panel. Number of participants in parentheses.

the children. We calculated the average normalized scores (i.e., the average of the scores obtained from the 52 tests) associated with cluster 1 (C_1 , in gray) and cluster 2 (C_2 , in black) from children belonging to the different age groups (ranges: 30–35, 36–41, 42–47, 48–53, 54–59, 60–65 and 66–71 months; Fig. 3a). We then calculated the probability that the children belonging to these two clusters attended either a public, private or no school at all (Fig. 3b). Notably, children belonging to C_2 (i.e., w. higher scores) were ~120% (all ages) more prevalent in private schools than children from C_1 (i.e., w. lower scores).

Next, we investigated if household conditions, like children living with both parents, mother only, or with mother and stepfather, influenced their classification into their respective clusters (Fig. 3c). Children from C_2 were ~43% more likely than those from C_1 to belong to a family with both parents. Similarly, we found that the average age of the parents (Fig. 3d) and their average educational level (Fig. 3e) influenced the classification groups, where older parents with a higher educational level were associated with C_2 (Supplementary Tables 2 and 3).

At first sight, the fact that children attending private schools were more likely to belong to C_2 could be explained by differences in the education programs provided by private vs. public schools. However, such a relationship could also be a consequence of other domestic factors such as parents' age or their educational level. We explored our database and found that parents of children attending private schools were older, on average, than those of children attending a public or no school (age father: $F = 50.75$, $P < 0.001$; age mother: $F = 58.51$, $P < 0.001$; Fig. 3f). Similarly, the average educational level of the parents of children attending private schools was higher than that of parents of children attending public or no school (years of schooling father: $F = 119.85$, $P < 0.001$; years of schooling mother: $F = 112.37$, $P < 0.001$; Fig. 3g). Therefore, the score differences in children attending public and private schools depended on factors existing outside the school²¹. Additionally, using a multivariable linear regression model, we explored how specific predictors (Pr_1 : type of school, Pr_2 : household, Pr_3 : father's age, Pr_4 : mother's age, Pr_5 : educational level of the father, Pr_6 : educational level of the mother) explained: (1) the global scores of the children (left panel in Fig. 3h), and (2) the classification clusters of the children (right panel in Fig. 3h). In Fig. 3h, significant coefficients for predictors (Pr_1 – Pr_6) compared against surrogates (obtained

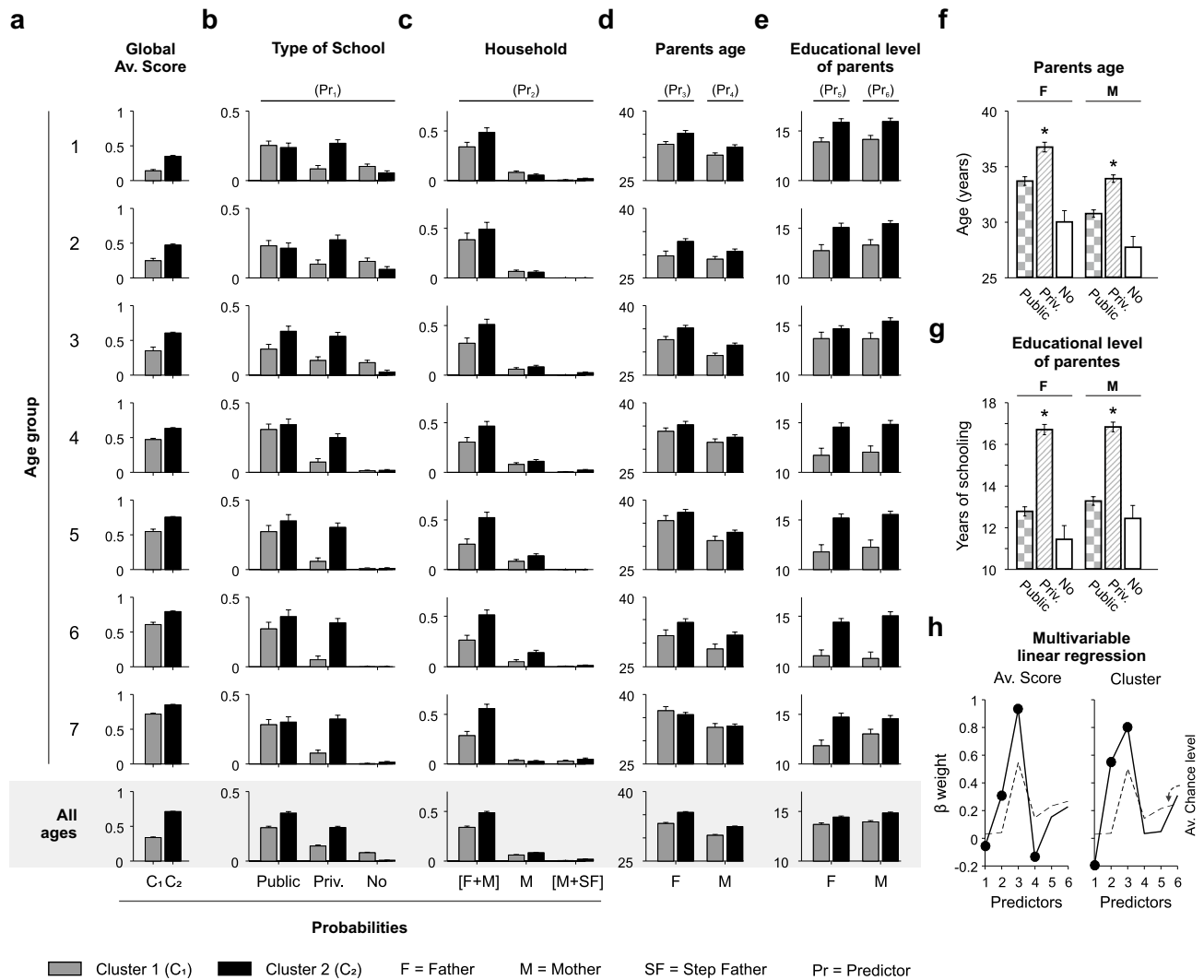


Figure 3. Influence of domestic and educational factors on cluster classification. (a) Global scores (i.e., the average from all tests) for the participants from cluster 1 (C_1 , in gray) and cluster 2 (C_2 , in black). (b) Probabilities that children from C_1 and C_2 went to public, private, or no-school (Predictor 1: Pr_1). (c) Probabilities that children from C_1 and C_2 lived with both parents, only with the mother or with mother and stepfather (Pr_2). (d) Average age of fathers (left, Pr_3) and mothers (right, Pr_4) of children from C_1 and C_2 . (e) Average years of schooling of fathers (left, Pr_5) and mothers (right, Pr_6) of children belonging to C_1 and C_2 . (f) Average age (f) and average years of schooling (g) of fathers (left) and mothers (right) of children attending public, private, or no-school. (h) Multivariate regression model exploring how the six predictors (Pr_{1-6}) influenced cluster assignment (significant regression weights marked with filled dots).

by permuting the predictors) are illustrated with filled dots. These results reveal that type of school, household, parents’ educational level, and parents’ age influence the scores and classification schemes of the children’s neuropsychological functions.

Integration of neuropsychological functions through development. Colormaps illustrating correlation matrices or neuropsychological functions help establish paired comparisons across tests (*v.gr.*, Fig. 1b). However, tests that concentrate high correlations cannot be easily identified using these graphs. Therefore, we implemented a ‘network representation’ where we distributed the 52 tests as discrete nodes along a circumference, with the lines’ thickness and colors connecting these dots representing the cross-correlation level between each pair of nodes (tests). These plots were made with data from each age group (Fig. 4a) and the entire dataset, including all ages (Fig. 4b; see Supplementary Table 1). Besides, we implemented centrality measures from network theory²² to explore each test’s relevance within this network of functions. Many of these metrics are based on the idea that well-interconnected nodes are likely to facilitate global, intermodular integration. We thus calculated the betweenness centrality (BC) of each node to identify relevant connections with other tests within the battery. We represented the BC of the tests as the area of the circular nodes in Fig. 4a,b. Interestingly, the ten tests with the highest BC exhibited a faster increase across age groups than the average of ten tests with the lowest BC (dotted line in Fig. 4c). This result suggests that expressive and receptive functions establish earlier their founda-

tions with abilities that promote the development of other skills. Interestingly, it is after 4 years of age that more complex abilities, such as math skills and cognitive flexibility, show complex connections.

The BC of the tests exhibited distinctive values (Fig. 5a) and consolidated across age groups (inset in Fig. 5b). To get an overview of the centrality for all ages, we calculated the summed BC for all tests across all age groups ($\Sigma(BC)$) and plotted the resulting data in linear (Fig. 5b) and logarithmic (Fig. 5c) scales. Using increasing thresholds given by different percentiles from the summed BC distribution, we created five groups of tests that exhibited increasing centrality values for all ages (Fig. 5c). Lower thresholds implied a more flexible rule with more tests having an equal or lower BC than the thresholded value (Fig. 5c). As a reference, the ten tests with the highest BC measures are listed in Fig. 4c.

Biometric systems serve to identify individuals by using physiological and behavioral predictors. We hypothesized that scores from tests with high BC could be employed to identify the children²³. We used the five groups of tests with different BC levels (in scales of gray) as predictors and combined them with supervised machine learning algorithms to predict the cluster to which the children belonged (Fig. 5d). We first cross-validated a k-nearest neighbor (KNN) classifier. We found that the classification improved with the number of tests employed as predictors and that ~20 neighbors optimized the classification procedure (black arrow in Fig. 5d). Next, we measured the performance of the trained classifier in identifying children from each cluster. Notably, identification of the best performers (i.e., children belonging to C_2) was more efficient and required fewer tests than for children from C_1 (right panel in Fig. 5e; KW-test, $P < 0.0001$ for all cases). Also, as expected, adding tests to the battery of predictors increased the classifier's prediction capacity when dealing with children from both C_1 and C_2 (right panel in Fig. 5e; RM One-Way ANOVA test, $P \geq 0.001$ for both cases).

Next, we used the same 20-neighbor KNN-classifier to challenge the prediction of the children's belonging to one of the four sub-clusters. We found similar tendencies as before: higher predictiveness required using fewer tests for higher performers (Fig. 5f). We confirmed these observations when applying the same KNN-classifier to Mexican children only (Fig. 5g). Altogether, these results illustrate how neuropsychological scores served to identify the children, where higher scores led to a more accurate classification. Therefore, neuropsychological tests' scores can be used as behavioral 'fingerprints' to identify participants with a high level of certainty.

Discussion

We explored age-related changes across neuropsychological functions from a large sample of young Latin-American children in the age range of 2.5–6 years (30–71 months of age). We used a battery of 52 standardized tests, designed for Spanish-speaking children¹³, to explore the developmental profile of receptive, expressive, attention/memory, processing, and executive functions.

We identified age-related gains in scores across all tasks, which reflect an organized development of neuropsychological functions, as has been reported^{4,24–28}. We applied a cluster analysis to our data and identified two main clusters (C_1 and C_2) associated with a specific configuration that suggests that not all functions developed equally in all children. Indeed, our results showed distinctive developmental trajectories of neuropsychological functions in preschoolers, as others have suggested³. Although similar age-related increases were observed for both cluster groups (i.e., similar slopes), C_1 lagged behind C_2 . Importantly, C_1 was not composed of learning disabled children since we controlled biological factors related to neurodevelopmental disorders, and it involved a large proportion of the entire sample (42%). However, C_1 children exhibited weaker performance in all tests, and particularly lower in left-hand recognition, instructions, and naming tests. These tests involve inter-modal information transfer, receptive language, and vocabulary, respectively. Differences in right/left hand recognition tests are usually linked to functional variabilities between right/left cerebral hemispheres. In sum, such cluster differences might be relevant to establish criteria to identify children with low performance in multiple neuropsychological domains and a potential developmental delay.

Nowadays, accurate assessment of neuropsychological functions in children requires collecting cognitive, linguistic, motor, executive, emotional, and behavioral information together with the social environment, school, and household conditions. Increasing evidence suggests that environmental factors influence neuropsychological functions. This influence is evident when considering children from unstable families and with disruptive family events²⁹, and early abilities can predict later educational achievements^{30–33}. Thus, human development has to be studied within a context³⁴, even if it's considered a 'homogeneous' one. Because young children tend to spend more time with their families, we explored such variables' influence on the preschoolers' neuropsychological development. In Latin-America, preschool attendance is not yet universal, and most of the children attend public preschools. Therefore, to reflect the characteristics of preschool attendance in a Spanish-language ambiance, we evaluated more children from public than from private schools, in addition to a small percentage of children that did not attend any preschool/day-care. The family composition and parents' educational level exhibited a wide dispersion. We confirmed that the children's scores and classification clusters were sensitive to these environmental variables, particularly those related to parents' education^{5,35–37}. Children belonging to the high-performance cluster (C_2) were more prevalent in private schools than children from the other cluster (C_1), and they were also more likely to belong to a family with both parents present. Public schools in Latin-America tend to have more children per classroom, with a single teacher giving classes from different subjects and a limited budget to acquire teaching material. However, both educational systems do also have plenty of similarities. For example, both types of schools have mentoring schemes that adjust the educational programs to the child's capacities. Similarly, they both inform and orient the parents regarding the child's special needs and provide a psychoeducational assessment. Public schools also have external education units where children with special needs are partially attended.

At first glance, one could be prone to think that the education provided by private schools in Latin American countries stimulates cognitive development to a higher degree than public schools do. Nonetheless, given that

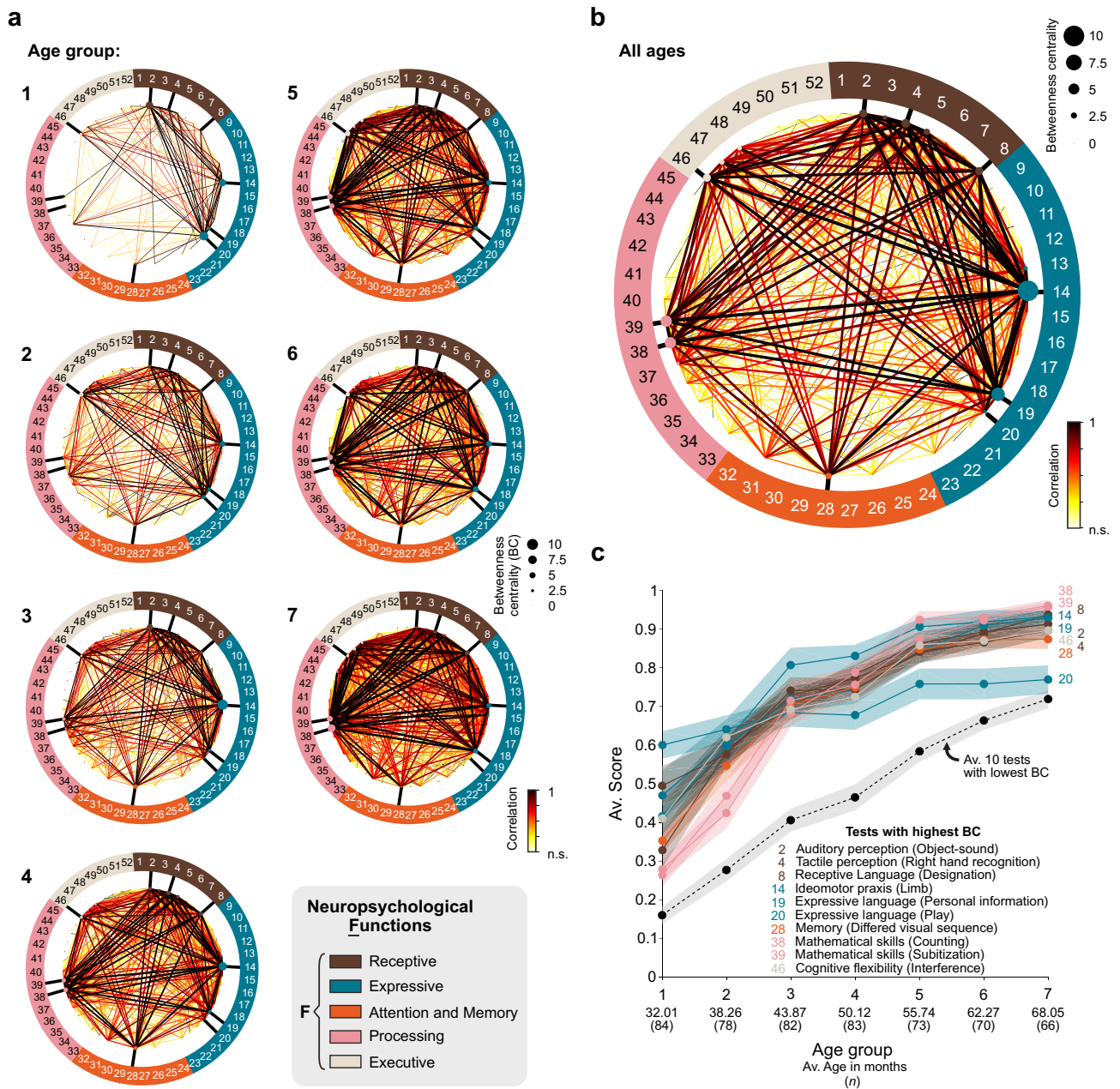


Figure 4. Consolidation/integration of critical neuropsychological functions through development. Network visualization of the cross-correlations across the 52 neuropsychological tasks using the scores from children belonging to each age group (a) or from all ages (b). Tasks appear as nodes radially arranged along a circumference covering the following domains: receptive (walnut brown), expressive (mosaic blue), attention/memory (orange), processing (flamingo pink), and executive (cloud cream-yellow) functions. The colors of the lines and their thickness represent the cross-correlation between pairs of tasks, whereas the area of each node is proportional to the betweenness centrality (BC) for each task. (c) Developmental profile for the ten tasks with the highest BC compared against the average of ten tasks with the lowest BC (dotted line).

many of our sample children just started school, this assumption might not be valid. We found that children from private schools were more likely to belong to a family with both parents present. Also, their parents had a higher level of education and were older than those of children attending public schools. Thus, distinctive familial contexts were associated with low- (C_1) and high- (C_2) performance clusters in several ways: family composition and parents' age and education level.

It has been reported that parent's level of education co-varies with the children's cognitive performance^{3,21,38,39}. In general, parents with high educational profiles tend to create enriched environments for their children⁴⁰⁻⁴², and use richer vocabulary^{43,44}, leading to faster language development, higher school attendance, and better performance in cognitive tests⁴⁵⁻⁴⁷. Other environmental factors might also contribute to children's development, such as financial resources, parental styles, the system of values, among others⁴⁸. Our results suggest that

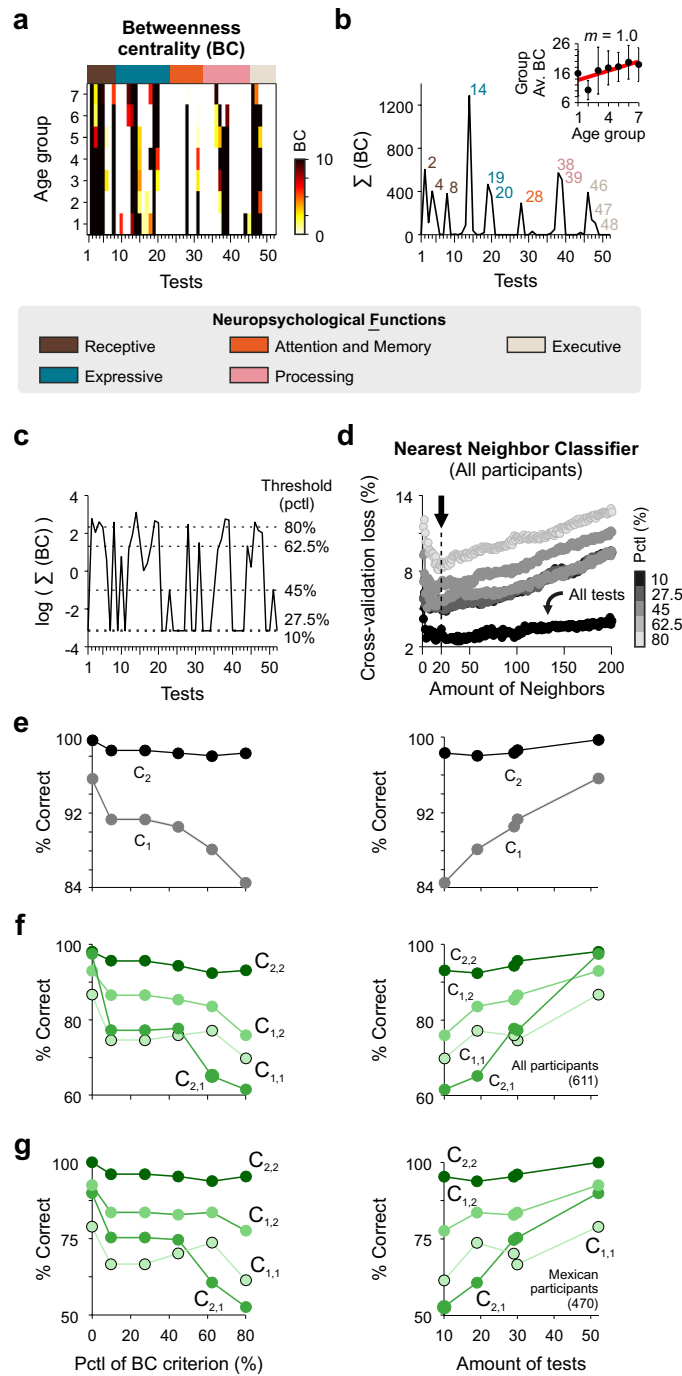


Figure 5. Supervised classification with centrality measures. **(a)** Colormap illustrating the betweenness centrality (BC) for each task as a function of age group. Binary BC was computed by thresholding each normalized correlation matrix against 0.5. Colorbar on the right. Summed BC (across ages) for the different tests in linear **(b)** and logarithmic **(c)** scales, respectively. **(d)** Cross-validation loss of a k-nearest neighbor (KNN) classifier as a function of the number of neighbors. Predictors consisted of the different tests associated with the different threshold values for BC. Optimal number of neighbors is around 20. **(e–g)** Performance of the KNN classifier as a function of the %BC criterion (left column), and as a function of the number of tests that correspond to that BC criterion (right column). The panels display the performance in classifying the children from the two main clusters for the entire dataset **(e)**, the four sub-clusters for the entire group **(f)**, and the four sub-clusters for the Mexican children only **(g)**. Number of participants in parentheses.

the higher number of C_2 children in private schools derives from more stable families, consolidated over more years, where parents have better jobs and, therefore, exhibit a higher socio-economical status, allowing them to afford a private education for their children. The family/school ambiance probably improves the children's development. Indeed, children's interaction with a variety of factors that promote or inhibit neuropsychological development starts well before they begin school²¹.

We found a general increase in performance over time as the children matured. Many of these improvements could be associated with a response to repeated experience and learning⁴⁹. However, we found that not all functions developed at the same time. To quantify the importance of individual 'network nodes' during this process, we implemented a developmental correlation analysis that included centrality measures²². We found that tests with the highest betweenness-centrality (BC) values were similar for all age groups and belonged to the five neuropsychological functions. The ten tests with the highest BC measures involved language (tests 8, 19, 20), auditory (test 2) and haptic (test 4) perception, praxis (test 14), visual memory (test 28), mathematical skills (tests 38, 39), and cognitive flexibility (test 46). In contrast, the ten tests with the lowest BC values did not involve processing and executive functions.

Strong correlation among tests might reflect that they share common underlying input variables. However, the lack of such correlations may not necessarily reflect independence between neuropsychological functions³. The interdependence among all these abilities is relevant because basic-level deficits (e.g., attention) could affect and limit the emergence of other abilities (e.g., language). Indeed, neuropsychological functions' development depends on frontal functioning orchestrated with other interacting systems' maturation, including attention, memory, language, and emotions¹². Our results show heterogeneous correlations among tasks, supporting a 'multidimensional' notion for neuropsychological functions. They also suggest that abilities with high centrality could help establish the foundations for other skills to emerge. Alternatively: high centrality abilities could require many other functions to consolidate. Both perspectives imply that brain regions interact and facilitate functional integration of more advanced skills.

Many researchers suggest that a sequence in the mastery of skills required for successful development through childhood, coinciding with the maturation of the cerebral cortex⁴. For example, previous works have shown how inhibitory control's development precedes maturation of other more elaborate functions, such as selective attention³. Our analysis revealed that, although functions were interrelated, their distinctive developmental trajectories were well separated already from the youngest age groups³. More (cumulative) changes probably occur during subsequent years. Some studies suggest a progression of executive skills in which proficiency is still not achieved by 12 years of age^{50,51}.

Neuropsychological functions are crucial to adjust and adapt to changing environmental demands. Delays in neuropsychological functions' maturation can affect how children interact with the environment, leading to other problems in cognitive, academic, and social dimensions^{4,52}. In this study, we controlled some aspects of the biological background, excluding from our sample children with genetic aberrations and early CNS insults that would skew their development. Still, we found clear distinctive developmental trajectories linked to domestic and educational factors. This finding supports the idea that the environment shapes the way internal (genetic) variables are expressed. Prenatal and postnatal experiences determined the child's development rendering a wide range of neuropsychological profiles evident as early as 30 months of age. Our approach highlights the importance of tracking developmental trajectories with an apparent heterogeneity across Latin-American countries.

Dynamic variations in the child-environment dyad give rise to distinctive phenotypes. Neurodevelopmental disorders tend to originate from basic-level deficits that have 'cascading' effects on other domains of the developing system, including cognition, executive functions, language, social functions, motor functions, and behavior control. Several diagnoses fall within these categories, including learning and intellectual disabilities, attention deficit hyperactivity disorder (ADHD), and communication disorders. Our results highlight the fact that identification and classification of the children with low scores requires a comprehensive battery of tests. Such instruments enable reliable diagnostics. Thus, an exciting area for future research will explore the implications of the deficits and developmental delays of neuropsychological functions associated with common neurodevelopmental disorders⁵³. It will be crucial to describe the differences at multiple levels (genetic, neural, cognitive, environmental) of typically developing children and children with developmental delays and how these differences increase/decrease over time¹.

Classification algorithms using large multidimensional biological datasets can be used for prognosis (predicting risk from neuropsychological markers), diagnosis, and therapeutic intervention. These methods can be useful for identifying low performing children and then testing the effectiveness of interventions. Proper classification would provide schools information about children's development and might help to diagnose possible neuropsychological disorders. Moreover, classification schemes could guide patient care and treatment of neurodevelopmental disorders, or brain-injured children with cognitive, linguistic, motor, executive, and behavioral therapies based on neuropsychological knowledge. In this context, one of our interests is to provide Latin-American schools, clinicians, and scientists with specific data to guide them when assessing preschoolers' development.

A common critique of neuropsychological tests is that their scores are imperfect indicators of latent variables. As mentioned before, neuropsychological functions are complex and multidimensional constructs and, therefore, interpretation of test scores is difficult to make³. Furthermore, the quantification of neuropsychological functions in children is more challenging than in adults. First, assessing these functions only constitutes a snapshot at a particular time. Second, the tests are limited and may not be sensitive to more subtle developmental changes. Another limitation of our work is that we did not perform a longitudinal study, meaning that we did not map the developmental trajectories for individual children. Longitudinal studies of children could provide information regarding the intra-individual rate and extent of development of such functions. Such approximations could

help classify individual differences. Furthermore, it would also be crucial to explore the degree of plasticity in the neural networks that mediate neuropsychological functions.

Instead of considering a single compromised neuropsychological function, this work explored and described a broader neuropsychological profile of preschoolers. Using clustering analysis at multiple levels, we assessed the children's performance for each task, the developmental stage of each function, and the interrelations between neuropsychological domains. We also explored the interplay between cognition and experience. Our results highlight the fact that not only biological and environmental factors drive and support cognitive development: there is a third factor related to how children build their knowledge, namely, how their own cognitive characteristics enable them to process the stimuli provided by their environment.

Methods

Participants. Data acquisition consisted of applying a Neuropsychological battery of tests (Supplementary Table 1) on a total of 643 typical (born at ≥ 37 weeks of gestation; birth weight between 2500 and 4000 g; no reports of prenatal, perinatal, or postnatal complications, or TBI, without a neurodevelopmental disorder diagnostic) Latin-American children (315 boys, 328 girls) ranging from 30 to 71 months of age. The database involved scores from 473 Mexican, 115 Colombian, and 55 Guatemalan children with similar socioeconomic backgrounds.

Tests. The evaluation instruments were based on procedures developed in cognitive neuroscience and were standardized for Spanish-speaking preschoolers¹³. The 52 subtests belong to twenty domains grouped into five classical cognitive functions^{5,54}: (1) *receptive functions* (domains: visual, auditory, haptic perception; they involve the processing of sensory information and receptive language), (2) *expressive functions* (domains: fine motor, ideomotor praxis, constructional abilities, graphic skills, gross motor, and expressive language), (3) *attention/memory functions* (domains: coding, delayed recall, attention), (4) *processing functions* (domains: reasoning, math, and pre-reading skills), and (5) *executive functions* (domains: cognitive flexibility/shifting, self regulation working memory and theory of mind). We used non-digital, manipulative material (pencil, paper, objects like dolls, cars, balls, among others; see Supplementary Table 1 for a detailed description of each test). These children were subdivided into seven well-balanced age groups, with the following number of children per group: 30–35 months: $n = 98$ (54 boys, 46 girls); 36–41 months: $n = 97$ (46 boys, 51 girls); 42–47 months: $n = 102$ (50 boys, 52 girls); 48–53 months: $n = 101$ (48 boys, 53 girls); 54–59 months: $n = 87$ (42 boys, 45 girls); 60–65 months: $n = 83$ (41 boys, 42 girls); 66–71 months: $n = 75$ (36 boys, 39 girls). We used a parent's clinical questionnaire, adapted from the *Evaluación Neuropsicológica Infantil*⁵⁵ to ascertain the presence of certain guidelines of typical development: (a) born at term (≥ 37 weeks of gestation)⁵⁶, (b) birth weight between 2500 and 4000 g (5.5–8.8 lbs.)⁵⁷, and (c) no reports of prenatal, perinatal or postnatal complications or trauma that could affect nervous system development^{54,58}. No children had any diagnosis of neurological or psychiatric disorders. Trained psychologists assessed the children and interviewed the parents in well-lit locations, free of distractions. Rooms were provided by the day-care or preschool centers, but the children who were not in day-care or preschool were evaluated in their homes. All testing was done individually, within a single week, and lasted approximately 2 h, divided into three sessions of 40–45 min. each. The order of tasks presentation was counterbalanced. We explained the whole procedure, the use, and confidentiality of data, and provided feedback to the parents of the children when requested. Examiners were sensitive to the regional differences in spoken Spanish when testing children from different regions/countries⁵⁹. Validation of these tests (consistency, reliability) was previously demonstrated^{60–62}. All children were monolingual native Spanish speakers. Agreements were reached with each public and private institution to assess the children and collect the data; then, parents signed informed consent before testing their children. The ethics committee of our institution approved the tests we performed in this study, which followed the principles of the Helsinki Declaration (#ET062009-62; Instituto de Neurociencias, Universidad de Guadalajara).

Analysis of neuropsychological functions. We normalized all the obtained scores to a scale 0–1 and grouped them into the five neuropsychological functions described above (receptive, expressive, attention/memory, processing, and executive functions). Due to our balanced group numbers, we did not consider size differences across age groups¹². We calculated the peak global cross-correlations across tasks using the scores from either the entire data set or from subjects belonging to each age group. We estimated the differences in average correlation by subtracting the average correlation values from contiguous age groups.

Cluster analysis. We partitioned the original data set into a discrete number of clusters by using the squared Euclidean distance and the k-means algorithm for cluster center initialization. We selected the optimal number of clusters by using the Calinski–Harabasz and the Davies–Bouldin criteria. We measured the absolute Euclidean distance between cluster centroids and the within-cluster sums of point-to-centroid distances as a function of age. From the original dataset, the scores from 32 children were partial and, therefore, could not be classified appropriately. We discarded these from subsequent analyses. We created cross-correlograms using data from subjects belonging to each cluster for each age group, and the average correlation difference was calculated between clusters for each age group. The average performance (and choice variance) from subjects belonging to each cluster for all ages or the different age groups was calculated and sorted either by neuropsychological function or by the summed correlations across tests (for the entire dataset). The polynomial coefficients (i.e., slopes and intercepts) from linear fits were obtained using a conventional least-squares algorithm.

We used measures from complex network analysis to characterize the relationship among tests²². In particular, we used a sensitive measure of global connectivity termed betweenness centrality (BC)^{63,64}. This binary

metric (using a cross-correlation threshold of 0.5) corresponds to the number of shortest paths that pass through each node. Consequently, ‘bridging’ nodes that connect separate parts of the network tend to have high BC. We calculated the summed BC across age groups and used a variable threshold in the summed BC to identify relevant tasks. Next, we used supervised machine learning tools to classify the children as belonging to each of the subclusters. The classification model was based on a k-nearest neighbor (KNN) classifier. We performed a tenfold cross validation and extracted the cross-validation loss of the predictive model. We also calculated the classification efficiency for the KNN-classifier as a function of the summed BC criterion described previously.

Analysis of contextual factors that influence neuropsychological classification. We explored whether domestic and educational environments of the children could predict their classification cluster. We used a multivariable linear regression model (MLRM) to quantify the regression coefficients for these predictors, as described before⁶⁵.

Statistical analysis. All statistical tests were performed using MATLAB R2016a (MathWorks, Inc.; Natick, USA). For group comparisons, we used Kruskal–Wallis multi-comparison tests and repeated measures ANOVA (RM-ANOVA) tests, with Bonferroni’s or Wilcoxon Signed Rank *posthoc* tests. We employed nonparametric tests whenever the assumptions required to use the parametric versions were not met. Normality tests were performed using Lilliefors, and Jarque–Bera tests⁶⁶. To quantify the significance of the MLRM coefficients, we contrasted the actual regression coefficients against 1000 surrogate data sets generated by shuffling the predictors, as previously described^{65,67}. With this approach, we established the empirical significance of the observed coefficients by comparing them against coefficients obtained with randomly permuted predictors (i.e., the surrogates). In other words, we tested the null hypothesis that the regression coefficients were generated by chance. We illustrate our group data as averages \pm S.E.M. with a significance set at $P \leq 0.05$. The number of children used for each analysis is represented inside parentheses in the corresponding figure panels.

Received: 31 July 2020; Accepted: 8 March 2021

Published online: 22 March 2021

References

1. D’Souza H, Karmiloff-Smith A. Neurodevelopmental disorders. *Wiley Interdiscip. Rev. Cognit. Sci.* **8**, 1398. <https://doi.org/10.1002/wcs>. (2017).
2. Grayson, D. S. & Fair, D. A. Development of large-scale functional networks from birth to adulthood: A guide to the neuroimaging literature. *Neuroimage* **160**, 15–31 (2017).
3. Klenberg, L., Korkman, M. & Lahti-Nuutila, P. Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Dev. Neuropsychol.* **20**, 407–428 (2001).
4. Anderson, V. Assessing executive functions in children: Biological, psychological, and developmental considerations. *Pediatr. Rehabil.* **4**, 119–136 (2001).
5. Lezak, M. D., Howieson, D. B., Loring, D. W. & Fischer, J. S. *Neuropsychological Assessment* 4th edn. (Oxford University Press, 2004).
6. Karmiloff-Smith, A. Nativism versus neuroconstructivism. In *Thinking Developmentally from Constructivism to Neuroconstructivism* (ed. Thomas, H.) (Routledge, 2018).
7. Denburg, N. L. *et al.* The orbitofrontal cortex, real-world decision making, and normal aging. *Ann. N. Y. Acad. Sci.* **1121**, 480–498 (2007).
8. Denburg, N. L., Tranel, D. & Bechara, A. The ability to decide advantageously declines prematurely in some normal older persons. *Neuropsychologia* **43**, 1099–1106 (2005).
9. Salthouse, T. A. *Major Issues in Cognitive Aging* (Oxford University Press, 2010).
10. Hooper, S. R. & Boyd, T. A. Neurodevelopmental learning disorders. In *Child Neuropsychology* (eds Obrzut, J. E. & Hynd, G. W.) (Academic Press, 1986).
11. Mar, R. A. The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia* **42**, 1414–1434 (2004).
12. Romine, C. B. & Reynolds, C. R. A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Appl. Neuropsychol.* **12**, 190–201 (2005).
13. Matute E, Roselli M, Beltrán-Navarro B, & Ardila, A. *Evaluación Neuropsicológica Infantil Preescolar*. Manual Moderno (2021).
14. Rosselli, M., Jurado, M. B. & Matute, E. Las Funciones Ejecutivas a través de la Vida *Neuropsicología*. *Neuropsiquiatría y Neurociencias* **8**, 23–46 (2008).
15. Arango-Tobón, O. E. *et al.* Relación entre lenguaje expresivo y receptivo y habilidades prelectoras. *Revista Latinoamericana de Psicología* **50**, 136–144 (2018).
16. Victora, C. G. *et al.* Maternal and child undernutrition: Consequences for adult health and human capital. *Lancet* **371**, 340–357 (2008).
17. World Health Organization. Use and interpretation of anthropometric indicators of nutritional status. *Bull. World Health Org.* **64**, 929–941 (1986).
18. Sattler, J. M. *Assessment of Children: Cognitive Foundations* 5th edn. (Sattler, 2008).
19. Sattler, J. M. *Assessment of Children: Behavioral, Social, and Clinical Foundations* 5th edn. (Sattler, 2008).
20. Gómez-Pérez, E., Ostrosky-Solís, F. & Prospéro-García, O. Desarrollo de la atención, la memoria y los procesos inhibitorios: Relación temporal con la maduración de la estructura y función cerebral. *RevNeurol* **37**, 0561–0567 (2003).
21. Ardila, A., Rosselli, M., Matute, E. & Guajardo, S. The influence of the parents’ educational level on the development of executive functions. *Dev. Neuropsychol.* **28**, 539–560 (2005).
22. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* **52**, 1059–1069 (2010).
23. Treviño, M., León, M. C. R. & Haro, B. Adaptive choice biases in mice and humans. *Front. Behav. Neurosci.* **14**, 99 (2020).
24. Korkman, M., Kemp, S. L. & Kirk, U. Effects of age on neurocognitive measures of children ages 5 to 12: A cross-sectional study on 800 children from the United States. *Dev. Neuropsychol.* **20**, 331–354 (2001).

25. Rosselli, M., Ardila, A., Bateman, J. R. & Guzman, M. Neuropsychological test scores, academic performance, and developmental disorders in Spanish-speaking children. *Dev. Neuropsychol.* **20**, 355–373 (2001).
26. Rosselli-Cock, M. *et al.* Evaluación Neuropsicológica Infantil (ENI): Una batería para la evaluación de niños entre 5 y 16 años de edad. Estudio normativo colombiano. *Revista de Neurología* **38**, 720–731 (2004).
27. Bauer, P. J. *et al.* NIH Toolbox Cognition Battery (CB): Measuring episodic memory. *Monogr. Soc. Res. Child Dev.* **78**, 34–48 (2013).
28. Howard, S. J. & Melhuish, E. An early years toolbox for assessing early executive function, language, self-regulation, and social development: validity, reliability, and preliminary norms. *J. Psychoeduc. Assess.* **35**, 255–275 (2017).
29. Sturge-Apple, M. L., Davies, P. T., Cicchetti, D., Hentges, R. F. & Coe, J. L. Family instability and children's effortful control in the context of poverty: Sometimes a bird in the hand is worth two in the bush. *Dev. Psychopathol.* **29**, 685–696 (2017).
30. Duncan, G. J. *et al.* School readiness and later achievement. *Dev. Psychol.* **43**, 1428 (2007).
31. Jordan, N. C. & Levine, S. C. Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Dev. Disabil. Res. Rev.* **15**, 60–68 (2009).
32. Romano, E., Babchishin, L., Pagani, L. S. & Kohen, D. School readiness and later achievement: replication and extension using a nationwide Canadian survey. *Dev. Psychol.* **46**, 995 (2010).
33. Crookston, B. T., Forste, R., McClellan, C., Georgiadis, A. & Heaton, T. B. Factors associated with cognitive achievement in late childhood and adolescence: The Young Lives cohort study of children in Ethiopia, India, Peru, and Vietnam. *BMC Pediatr.* **14**, 253 (2014).
34. Budwig, N., Turiel, E. & Zelazo, P. *New Perspectives on Human Development* (Cambridge University Press, 2017).
35. Ardila, A. Directions of research in cross-cultural neuropsychology. *J. Clin. Exp. Neuropsychol.* **17**, 143–150 (1995).
36. Fletcher-Janzen, E., Strickland, T. L. & Reynolds, C. R. *The Handbook of Cross-Cultural Neuropsychology* (Plenum, 2000).
37. Rosselli, M. & Ardila, A. The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain Cogn.* **52**, 326–333 (2003).
38. Anderson, V., Lajoie, G. & Bell, R. *Neuropsychological Assessment of the School-Aged Child* (University of Melbourne, 1995).
39. Matute-Villaseñor, E., Sanz-Martín, A., Gumá-Díaz, E., Rosselli, M. & Ardila, A. Influencia del nivel educativo de los padres, el tipo de escuela y el sexo en el desarrollo de la atención y la memoria. *Revista latinoamericana de psicología* **41**, 257–276 (2009).
40. Hoff, E. The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Dev.* **74**, 1368–1378 (2003).
41. Hoff, E. Causes and consequences of SES-related differences in parent-to-child speech. In *Socioeconomic Status, Parenting, and Child Development* (eds In, M. *et al.*) (Lawrence Erlbaum Associates, Inc, 2003).
42. Beyens, I., Valkenburg, P. M. & Piotrowski, J. T. Developmental trajectories of parental mediation across early and middle childhood. *Hum. Commun. Res.* **45**, 226–250 (2018).
43. Hoff, E., Laursen, B. & Tardif, T. Socioeconomic status and parenting. In *Handbook of Parenting* (ed. Bornstein, M. H.) (Erlbaum, 2002).
44. Hoff-Ginsberg, E. Mother-child conversation in different social classes and communicative settings. *Child Dev.* **62**, 782–796 (1991).
45. Ganzach, Y. Parents' education, cognitive ability, educational expectations and educational attainment: Interactive effects. *Br. J. Educ. Psychol.* **70**, 419–441 (2000).
46. Teachman, J. D. Family background, educational resources, and educational attainment. *Am. Sociol. Rev.* **52**, 548–557 (1987).
47. White, K. R. The relation between socioeconomic status and academic achievement. *Psychol. Bull.* **91**, 461–481 (1982).
48. Ayala S. La deserción escolar: ¿ un callejón sin salida? In: *Los Retos en la Educación del Siglo XXI* (ed. Matute, E., & Romo, R. M.), 77–101 (2001).
49. Restak, R. M. *The Secret Life of the Brain* (Joseph Henry Press/National Academies Press, 2001).
50. Becker, M. G., Isaac, W. & Hynd, G. W. Neuropsychological development of nonverbal behaviors attributed to “frontal lobe” functioning. *Dev. Neuropsychol.* **3**, 275–298 (1987).
51. Passler, M. A., Isaac, W. & Hynd, G. W. Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Dev. Neuropsychol.* **1**, 349–370 (1985).
52. Ewing-Cobbs, L. & Fletcher, J. M. Neuropsychological assessment of head injury in children. *J. Learn Disabil.* **20**, 526–535 (1987).
53. Astle, D. E., Bathelt, J., Team, C. & Holmes, J. Remapping the cognitive and neural profiles of children who struggle at school. *Dev. Sci.* **22**, e12747 (2019).
54. Aylward, G. P. *Infant and Early Childhood Neuropsychology* (Plenum Press, 1997).
55. Matute, E., Rosselli, M., Ardila, A. & Ostrosky-Solís, F. *Evaluación Neuropsicológica Infantil* (Manual Moderno, 2007).
56. March of Dimes, PMNCH, Save the Children, & World Health Organization. Born too Soon: The Global Action Report on Preterm Birth (World Health Organization) (2012).
57. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr.* **450**, 76–85 (2006).
58. Reynolds, C. R., & Fletcher-Hanzen, E. *Handbook of Clinical Child Neuropsychology*. Springer (2010).
59. Siedlecki, K. L. *et al.* Do neuropsychological tests have the same meaning in Spanish speakers as they do in English speakers?. *Neuropsychology* **24**, 402–411 (2010).
60. Beltrán-Navarro, B., Abreu-Mendoza, R., Matute, E. & Rosselli, M. Development of early numerical abilities of Spanish-speaking Mexican preschoolers: A new assessment tool. *App. Neuropsych. Chil.* **7**, 117–128 (2018).
61. Beltrán-Navarro, B. G.-R. A., Matute, E. & Rosselli, M. El desarrollo de habilidades de conciencia fonológica de niños mexicanos 2 años 6 meses a 5 años 11 meses de edad. *EduPsykhé* **15**, 31–51 (2016).
62. Beltrán-Navarro, B. M. E. & Rosselli, M. El desarrollo de habilidades de memoria de niños mexicanos de 3 años a 4 años 11 meses de edad. *Rev. Neuropsic. Neuropsiq. Neuroc.* **18**, 35–56 (2018).
63. Brandes, U. A faster algorithm for betweenness centrality. *J. Math. Sociol.* **25**, 163–177 (2001).
64. Kintali S. Betweenness centrality: Algorithms and lower bounds. [arXiv:0809.1906](https://arxiv.org/abs/0809.1906) (2008).
65. Chamorro, Y., Trevino, M. & Matute, E. Educational and cognitive predictors of pro- and antisaccadic performance. *Front. Psychol.* **8**, 2009 (2017).
66. Conover, W. J. *Practical Nonparametric Statistics* (Wiley, 1980).
67. Trevino, M., Castiello, S., Arias-Carrion, O., De la Torre-Valdovinos, B. & Coss, Y. L. R. M. Isomorphic decisional biases across perceptual tasks. *PLoS ONE* **16**, e0245890 (2021).

Acknowledgements

This work was partially supported by CONACyT Ciencia Básica 2015-01 #251760, and by a PROMEP grant (UDGPTC1217) awarded to BBN. RM-C received a M.Sc. scholarship from CONACyT. We gratefully acknowledge Dr. Mónica Rosselli and Dr. Alfredo Ardila for their invaluable contribution in all stages of the design, data collection and standardization of the ENI-P Battery. Thanks to Liliana Castillejo (Ph.D), Claudia Delfín (Ph.D), Paola López García (Msc), María Elena Navarro Calvillo (Ph.D), Yunque Zalapa Medina (Ph.D) and Hugo Guadalupe Canto Pech (Ph.D) who supervised the data collection from Mexican children. Thanks to Nadyezhda Van Tuylen (Msc) who supervised the Guatemalan team, whereas Olber Eduardo Arango Tobón (Msc), and

Martha Lucia Miranda Giraldo (Msc) coordinated the work in Colombia. Additionally, we would like to thank the institutions, children and families who contributed and participated in this study. We are grateful for the constructive feedback from the reviewers.

Author contributions

M.T.: conceived and developed analysis algorithms, analyzed data, made figures, drafted and wrote the manuscript. B.B.N.: participated in the design of the neuropsychological battery, coordinated the data collection work, the standardization, reliability, and validity analyses of the tests, and contributed partially to writing the manuscript. R.M.: edited figures and tables. E.M.: conceived and directed all research, participated actively in all aspects of the work and contributed directly to the manuscript. All authors approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-85891-2>.

Correspondence and requests for materials should be addressed to M.T. or E.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021, corrected publication 2021