

# Métodos de clusterización para identificar diversidad cognitiva en niños y niñas

Implementación de métodos de clusterización para la identificación de diversidad cognitiva en dos muestras de niños y niñas provenientes del Área Metropolitana de Buenos Aires.

Giovannetti, Federico<sup>a, b</sup> Pietto, Marcos

Luis<sup>a, c</sup> Lipina, Sebastián Javier<sup>a</sup>

Kamienskowski, Juan Esteban<sup>d</sup> Segretin, María Soledad<sup>a</sup>

<sup>a</sup> Unidad de Neurobiología Aplicada (UNA, CEMIC-CONICET)

<sup>b</sup> Instituto de Cálculo (IC, FCEyN- CONICET)

<sup>c</sup> Universidad Argentina de la Empresa (UADE)

<sup>d</sup> Laboratorio de Inteligencia Artificial Aplicada (LIAA, FCEyN- CONICET)

## Antecedentes

Las ciencias del desarrollo reconocen que las diferencias individuales en el desarrollo cognitivo pueden ser adaptaciones al contexto y no necesariamente desempeños atípicos. Los análisis de clústeres ayudan a identificar perfiles diversos, pero faltan estudios que integren datos de distintas evaluaciones y poblaciones, lo que limita su uso en el diseño de intervenciones.

## Objetivos

A. Analizar la diversidad de perfiles de desempeño en tareas con demandas ejecutivas en niños y niñas de edad preescolar y de bajo nivel socioeconómico del Área Metropolitana de Buenos Aires.

B. Comparar la implementación de técnicas de clusterización en dos muestras con diferente composición demográfica.

## Métodos

Se analizaron dos muestras de proyectos desarrollados en la UNA: la muestra A estuvo compuesta por 58 participantes (Medad=5.82; DS=0.28), y la muestra B por 106 participantes (Medad=5.34; DS=0.33). Para ambas muestras, se evaluaron procesos de control inhibitorio (CI), memoria de trabajo (MT) y planificación (PL).

Los datos del desempeño fueron analizados mediante distintos métodos de clusterización (k-means, PAM, Diana y jerárquicos) y distancias (euclídea y manhattan), iterando en distinta cantidad de clústeres (k) para cada muestra. Posteriormente se analizaron y compararon las soluciones obtenidas para cada muestra.

Todos los análisis se realizaron mediante el lenguaje R versión 4.4.2 ([R Core Team 2024](#)), en Rstudio ([R-studio?](#)). Los principales paquetes utilizados fueron dplyr ([Wickham et al. 2023](#)), ggplot2 ([Wickham et al. 2025](#)) y diceR ([Chiu and Talhouk 2025](#)).

## Resultados

Se seleccionó un número final de k=3 mediante K-medias y PAM. Los grupos difirieron significativamente entre sí en todas las tareas ( $p < .05$ ), mostrando similitudes y diferencias para cada muestra.

- **Clúster 1** presentó desempeños y tiempos de reacción (TR) bajos en CI en ambas muestras. Para MT, mostró desempeños bajos en la muestra A y altos en la muestra B.
- **Clúster 2** presentó desempeños medios en CI en ambas muestras, con RT altos en la muestra A, acompañados de desempeños bajos en PL.
- **Clúster 3** tuvo desempeños medios-altos en CI y PL en ambas muestras. En los TR de CI, presentó TR bajos en la muestra A, y altos en la muestra B.

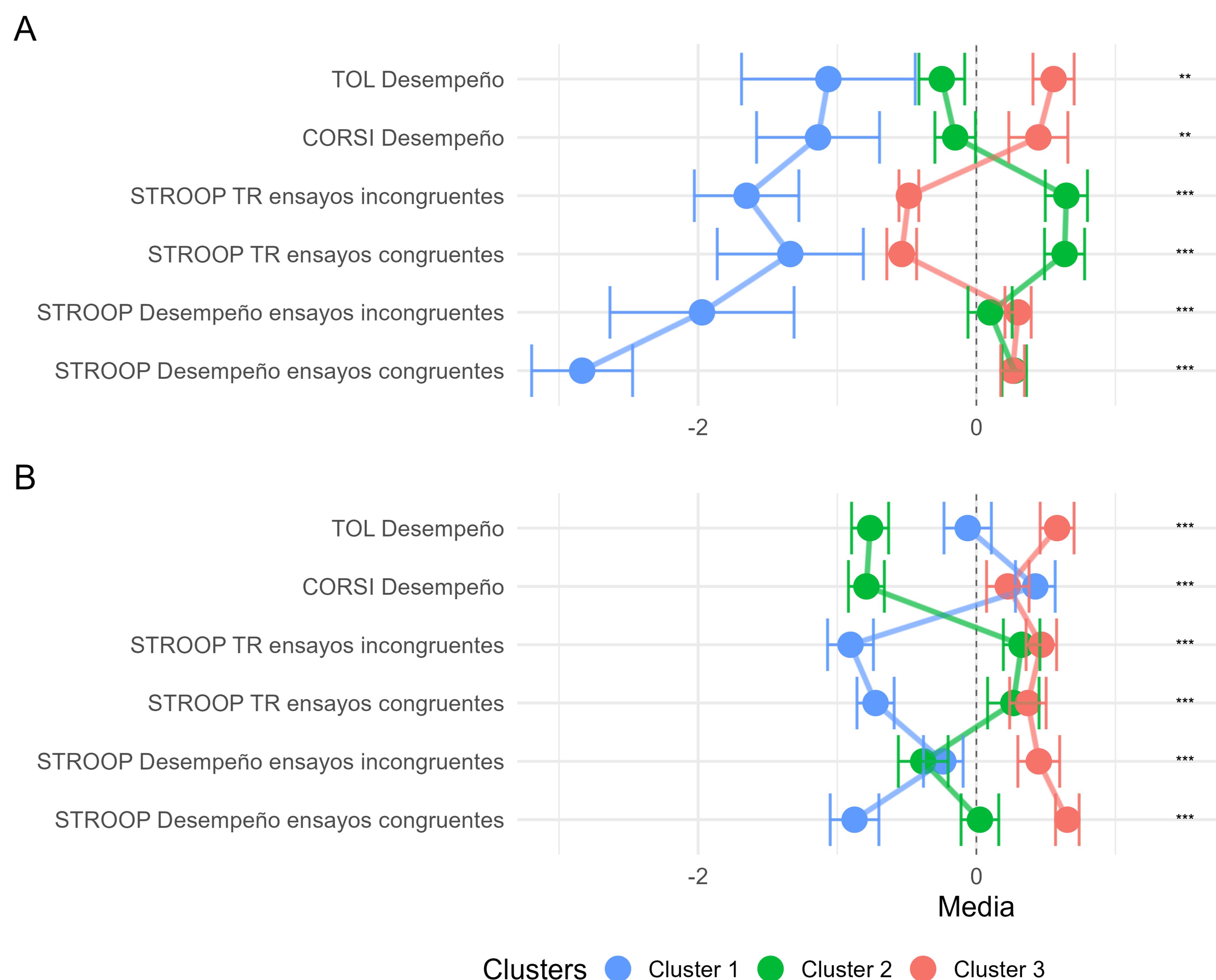


Figura 1. Comparación los desempeños y tiempos de reacción de los clústeres generados para cada muestra (valores z).\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .



Figura 2. Imágenes de las inmediaciones del barrio donde habitaban los/as participantes de las muestras A y B respectivamente.

## Conclusiones

Los métodos de agrupamiento permitieron identificar perfiles heterogéneos entre y dentro de las muestras. Estos resultados contribuyen al desarrollo de un enfoque metodológico relevante en el contexto de intervenciones que contemplen la diversidad cognitiva infantil.

## Referencias

- Chiu, Derek, and Aline Talhouk. 2025. *diceR: Diverse Cluster Ensemble in r*. <https://github.com/AlineTalhouk/diceR/>.  
R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.  
Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, Dewey Dunnington, and Teun van den Brand. 2025. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://ggplot2.tidyverse.org>.  
Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org>.

