

ASOFAMECH

Metodology and results.

Statistical Analysis.

Statistical analysis was performed with R (Version 4.2.3) and Rstudio (Version 2023.03.0+386) statistical softwares. All data were analyzed using one-tailed tests with a significance level of $\alpha = 0.05$. Confidence intervals (CI) were constructed with a confidence level of $1 - \alpha = 0.95$, using the t Student distribution.

Intra class correlation for evaluators.

Intra-Class Correlation (ICC) is a measure used in statistics to determine how strongly units in the same group resemble each other. It's often used in reliability studies where you want to assess the consistency or homogeneity within groups.

Negative ICC estimates are possible and can be interpreted as indicating that the true ICC is low, that is, two members chosen randomly from any class vary almost as much as any two randomly chosen members of the whole population.

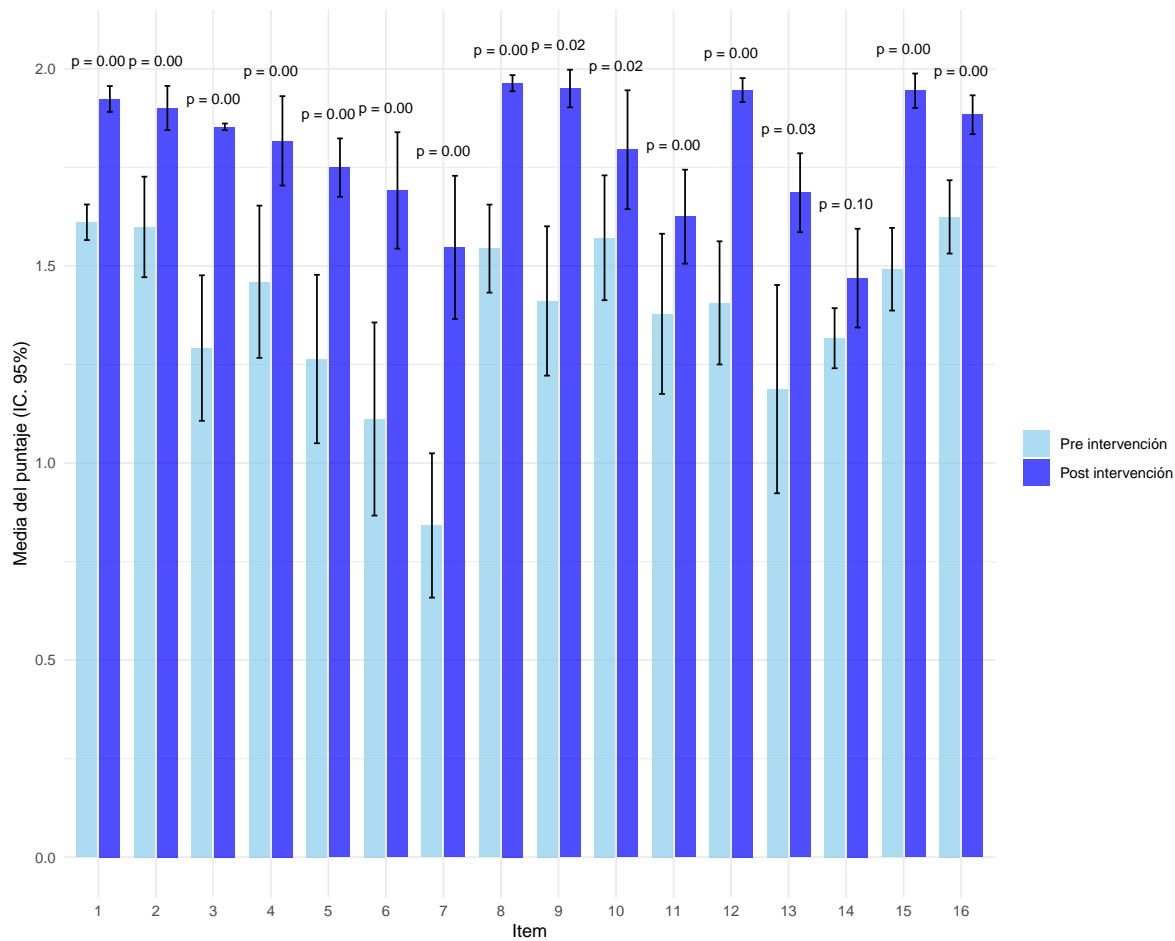
Mean comparison between pre- and post-intervention,

For the mean comparison between pre- and post-interventions for each item score and total item scores, the Wilcoxon signed-rank test was used, which is a non-parametric alternative to the paired t-test. This test was used when the assumptions of the paired t-test (such as the normal distribution of differences) were not met. It is appropriate to compare two related

| Grupo | item1_time1_icc_value | item1_time1_p_value | item1_time2_icc_value | item1_time2_p_value |
|-------|-----------------------|---------------------|-----------------------|---------------------|
| 1 | -0.17 | 1 | NaN | NaN |
| 2 | -0.07 | 1 | -0.07 | 1 |
| 3 | -0.08 | 1 | -0.08 | 1 |
| 4 | -0.09 | 1 | -0.09 | 1 |

samples when the data are ordinal or when interval-level measurements do not follow a normal distribution.

Scores for each item.



Total item scores.

