

Fleiss Multi-Rater Kappa

Test: Fleiss Multi-Rater Kappa

Design: Agreement among G1, G2, and G3 treated as three exchangeable raters of three grade categories.

Variables: categorized G1, categorized G2, categorized G3

Null hypothesis: Average observed multi-rater agreement equals chance agreement from pooled category margins.

Formula: $\kappa_F = (P_{\text{bar}} - P_e) / (1 - P_e)$, $P_i = [\sum_j n_{ij}^2 - k] / [k(k-1)]$.

Source rows: 649

Calculated metrics:

fleiss_kappa: 0.6421256617091

observed_agreement: 0.784283513097072

expected_agreement: 0.397228401641971

raters: 3

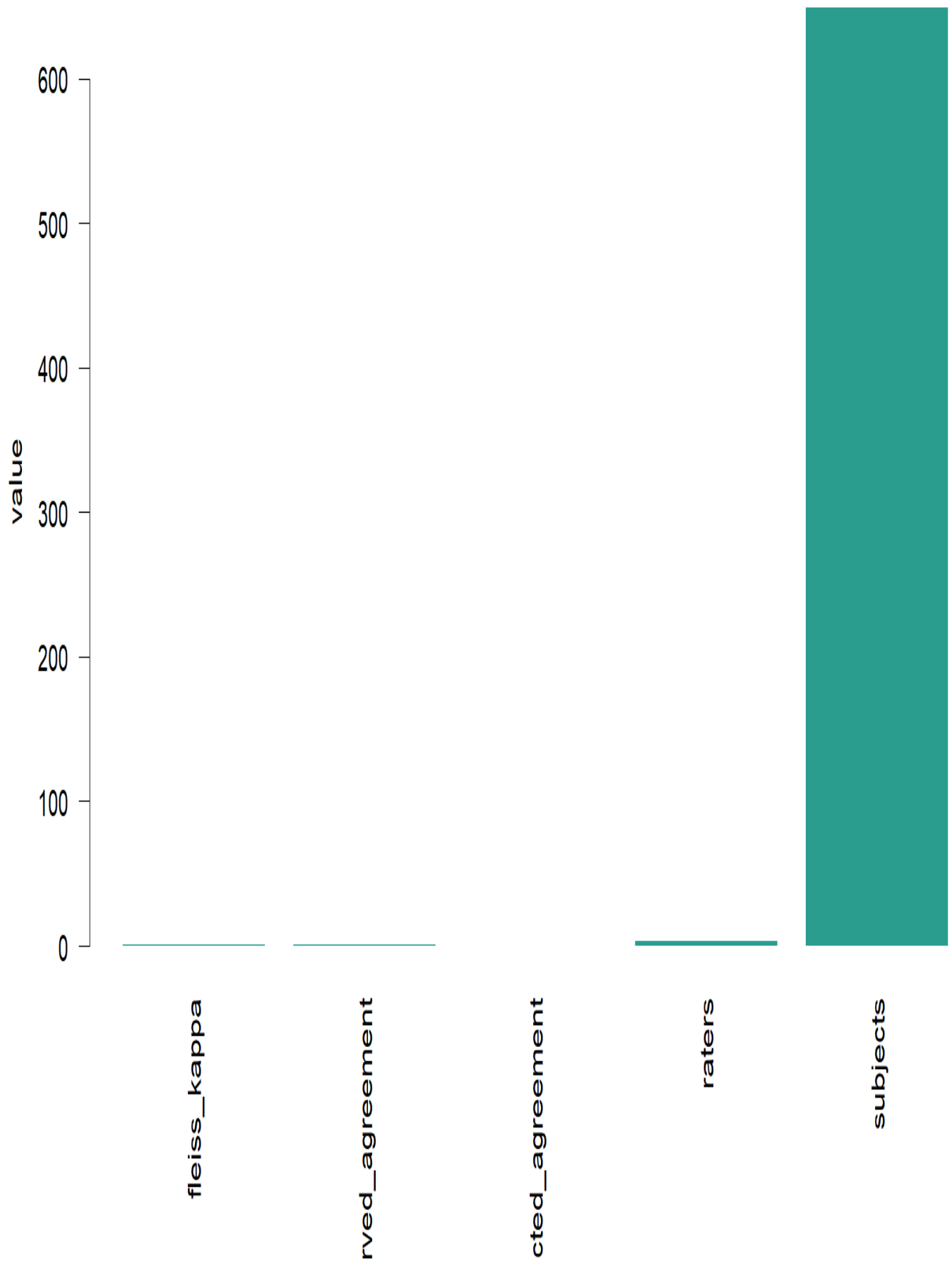
subjects: 649

Interpretation:

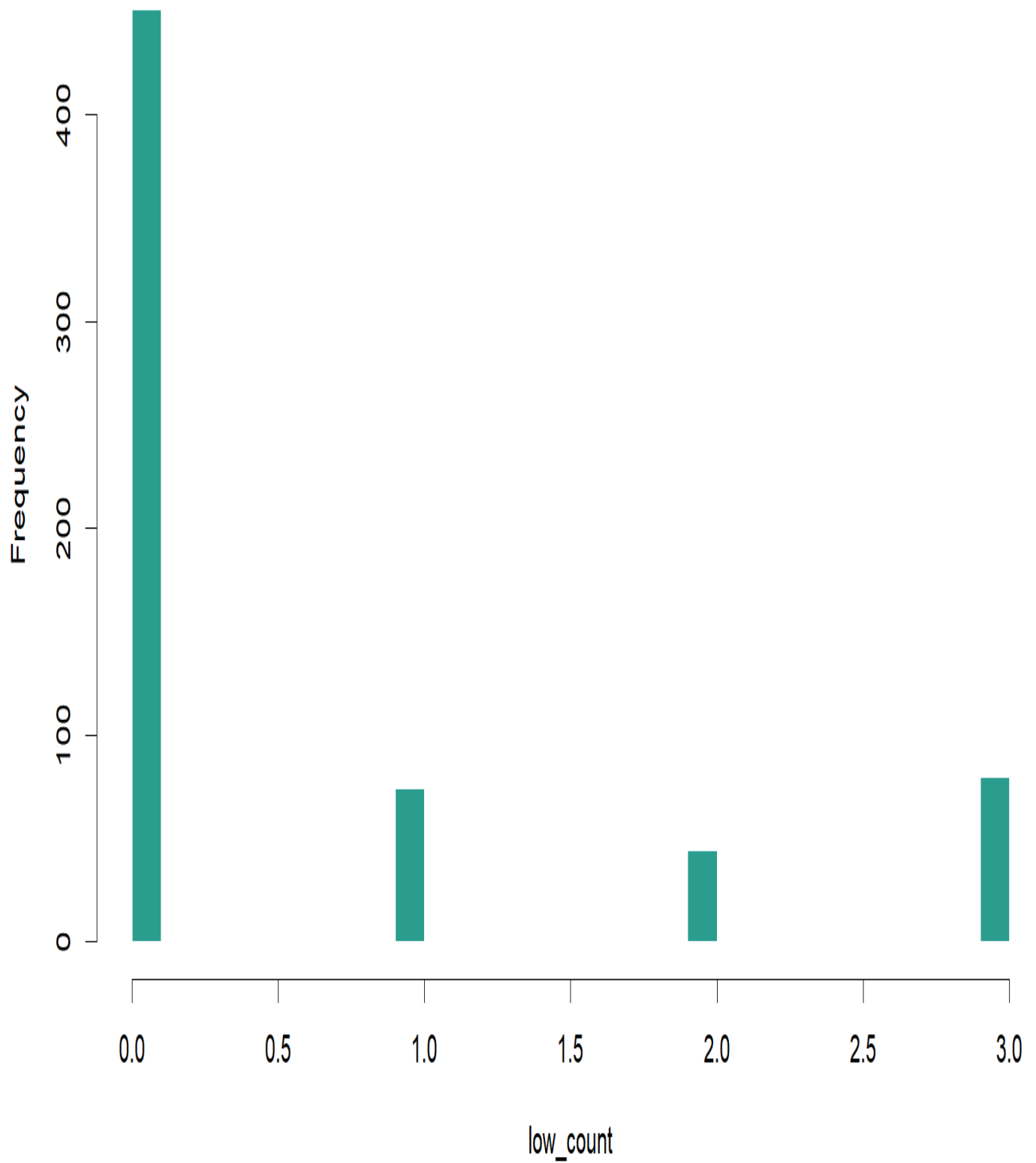
Pooled margins define chance agreement across three raters.

Fleiss kappa 0.6421256617091

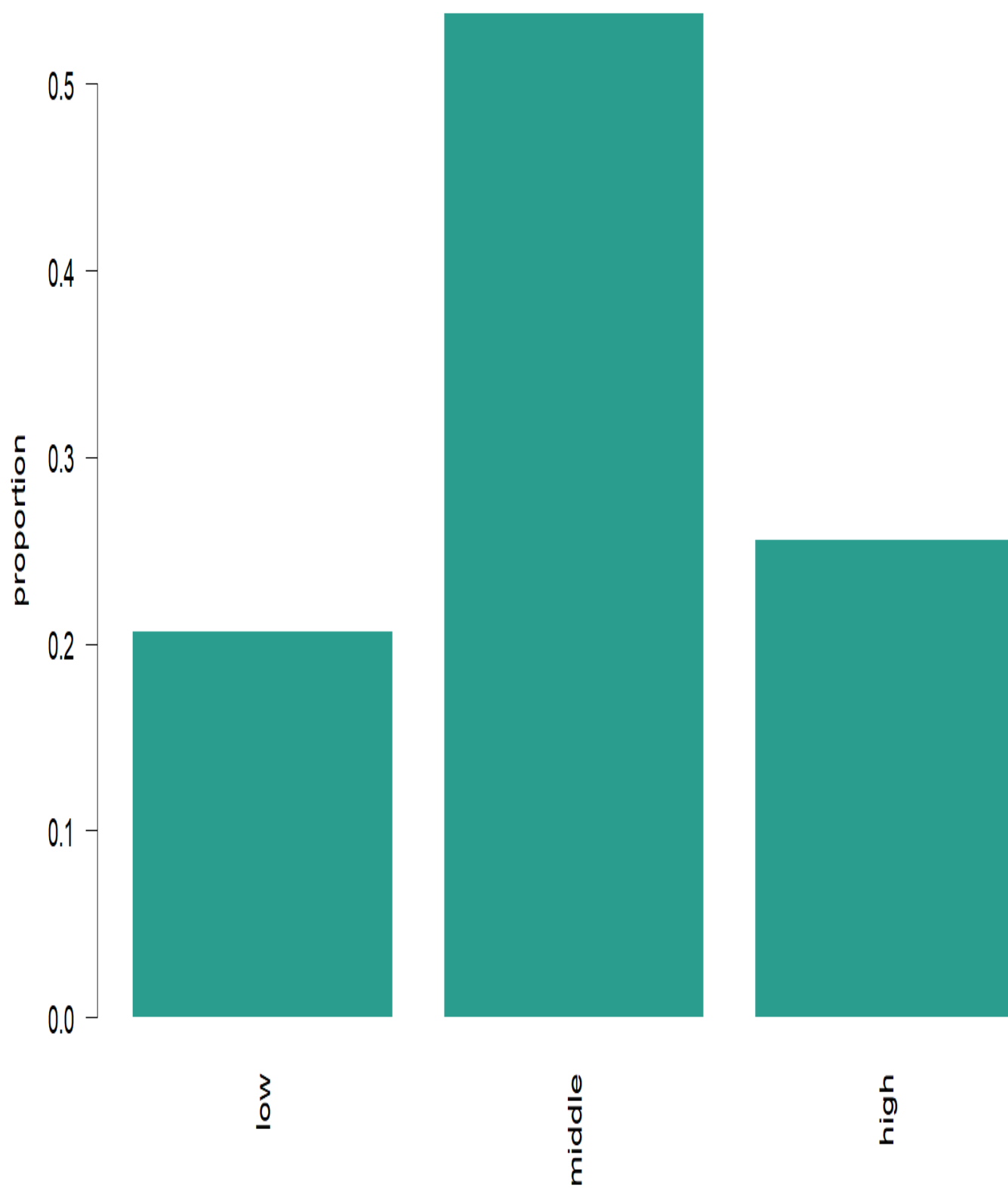
Fleiss Multi-Rater Kappa - primary metrics



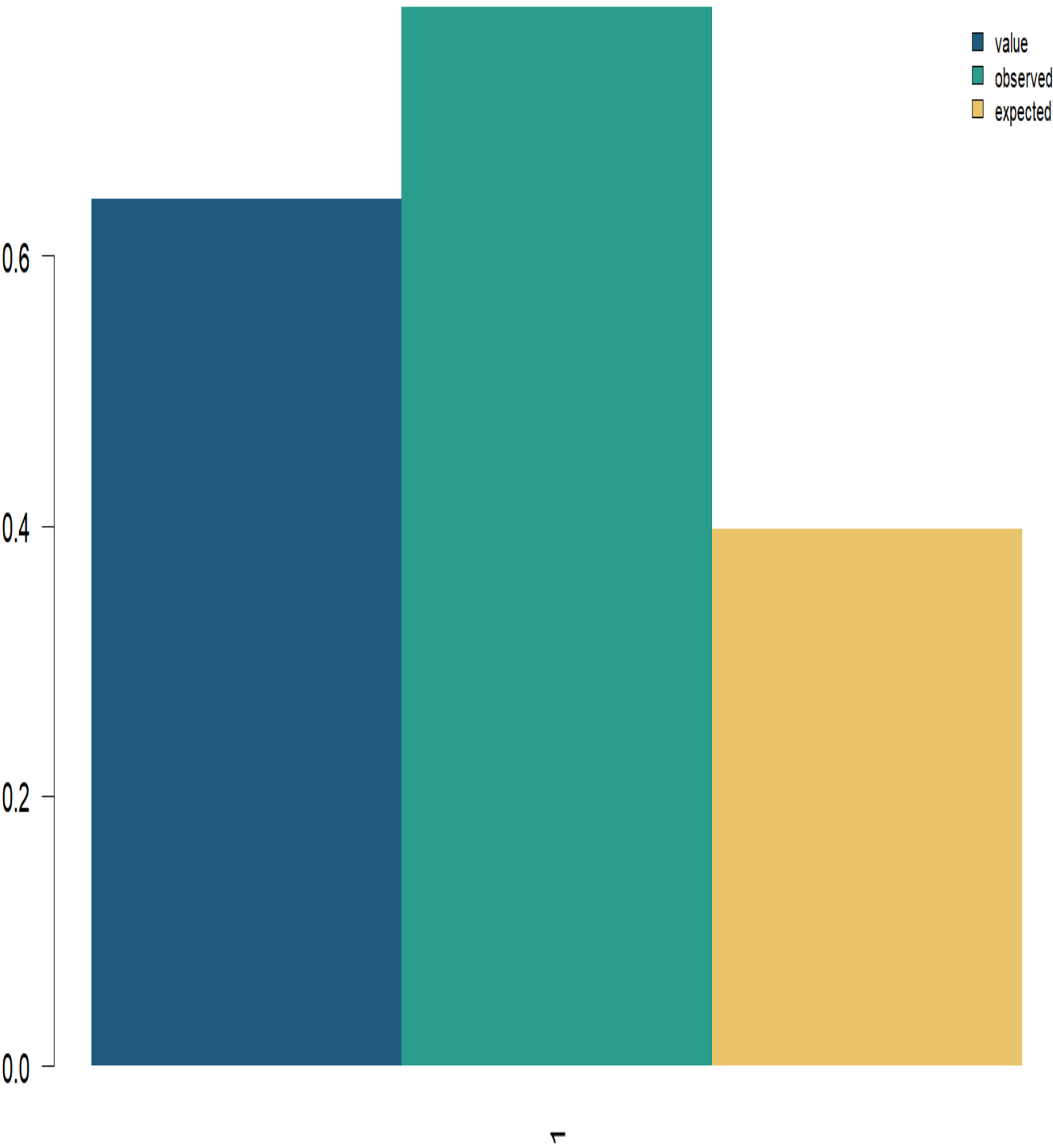
Fleiss Multi-Rater Kappa - subject category counts



Fleiss Multi-Rater Kappa - category proportions



Fleiss Multi-Rater Kappa - agreement components



Fleiss Multi-Rater Kappa verified results

