

Test-Retest Reliability

Standalone verified Python analysis

Test: Test-Retest Reliability

Design: Temporal stability of continuous G1 and G2 scores using Pearson association and change-score error.

Variables: G1, G2

Null hypothesis: The test and retest scores have zero linear stability.

Formula: $r_{tt} = \text{cor}(G1, G2)$; $SEM_{\text{change}} = SD(G2 - G1) / \sqrt{2}$.

Source rows: 649

Calculated metrics:

pearson_r: 0.8649816303085818

p_value: 6.3737947807967775e-196

mean_change: 0.17103235747303544

sd_change: 1.4792888099389643

standard_error_measurement: 1.0460151488412195

n_pairs: 649

Independent reference values:

pearson_r: 0.8649816303085818

p_value: 6.3737947807967775e-196

mean_change: 0.17103235747303544

sd_change: 1.4792888099389643

standard_error_measurement: 1.0460151488412195

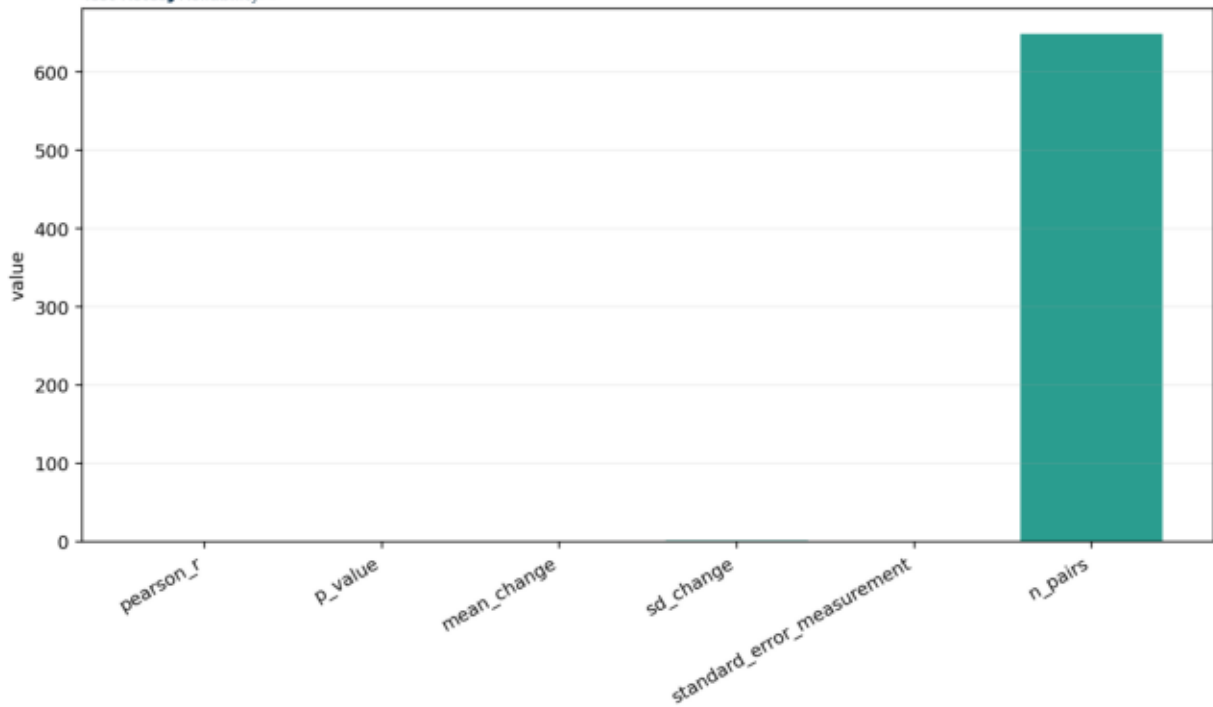
n_pairs: 649

Interpretation:

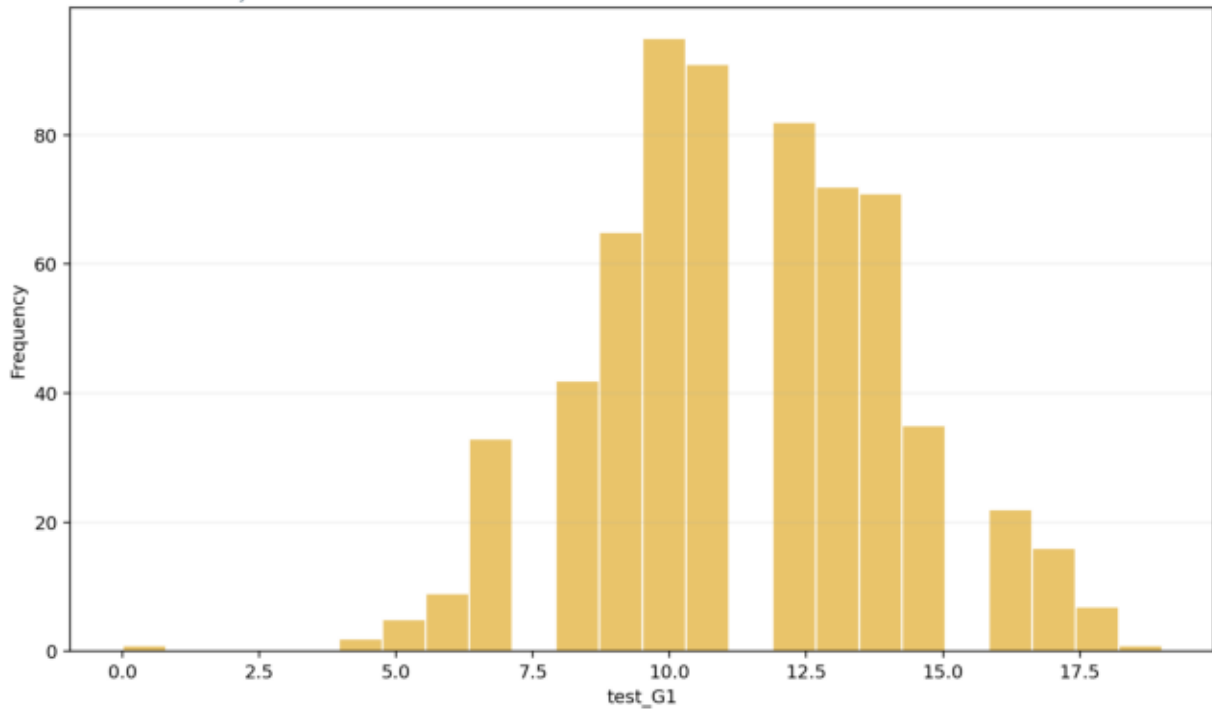
Rank-order stability and systematic score change are reported separately.

Test-retest $r = 0.864982$; mean change = 0.171032.

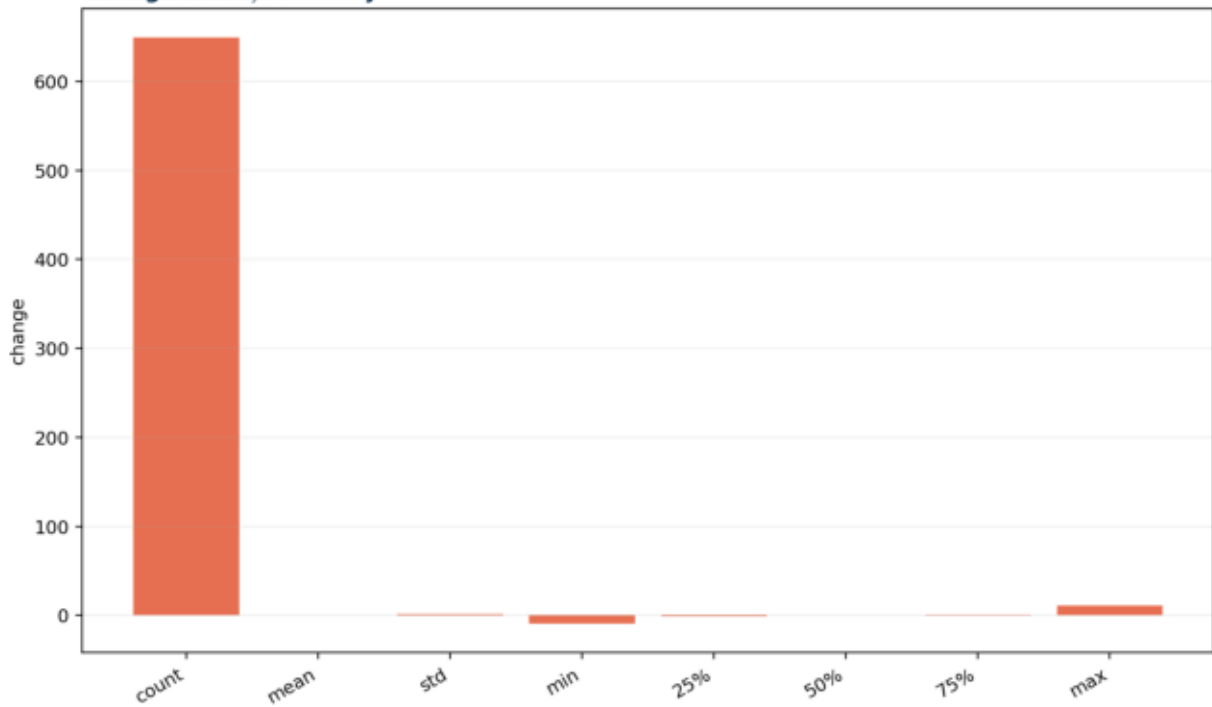
Primary Metrics



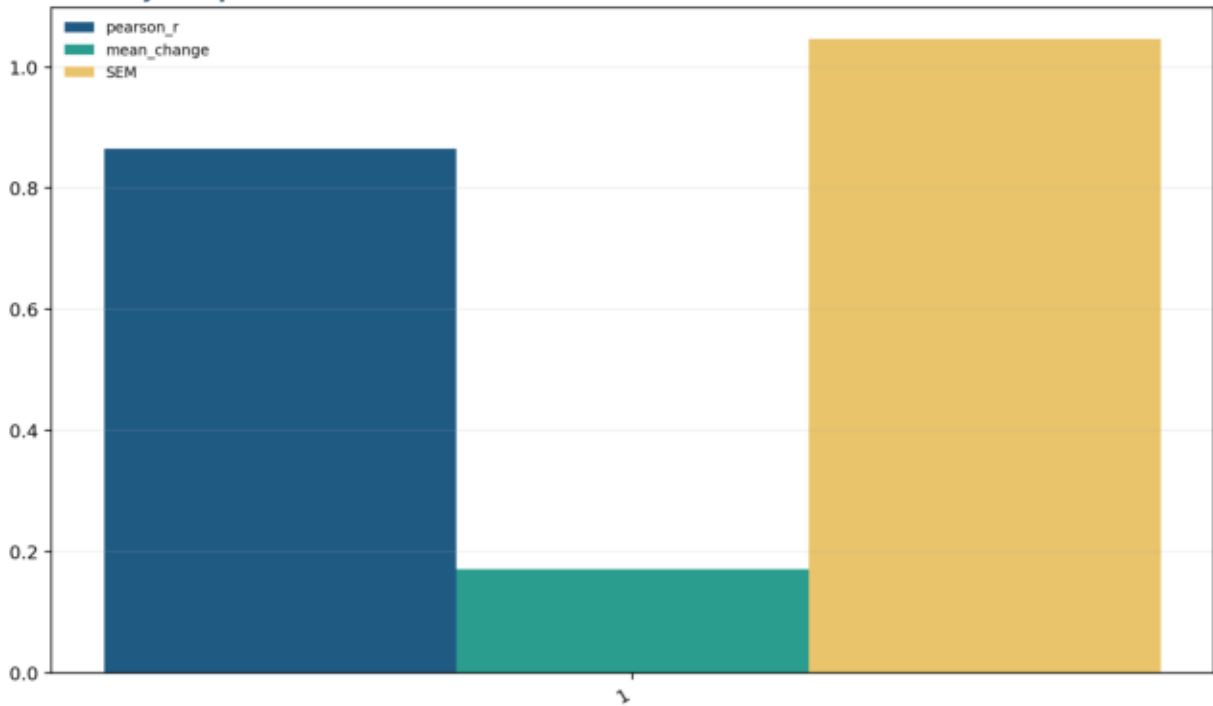
Paired Test Retest Scores



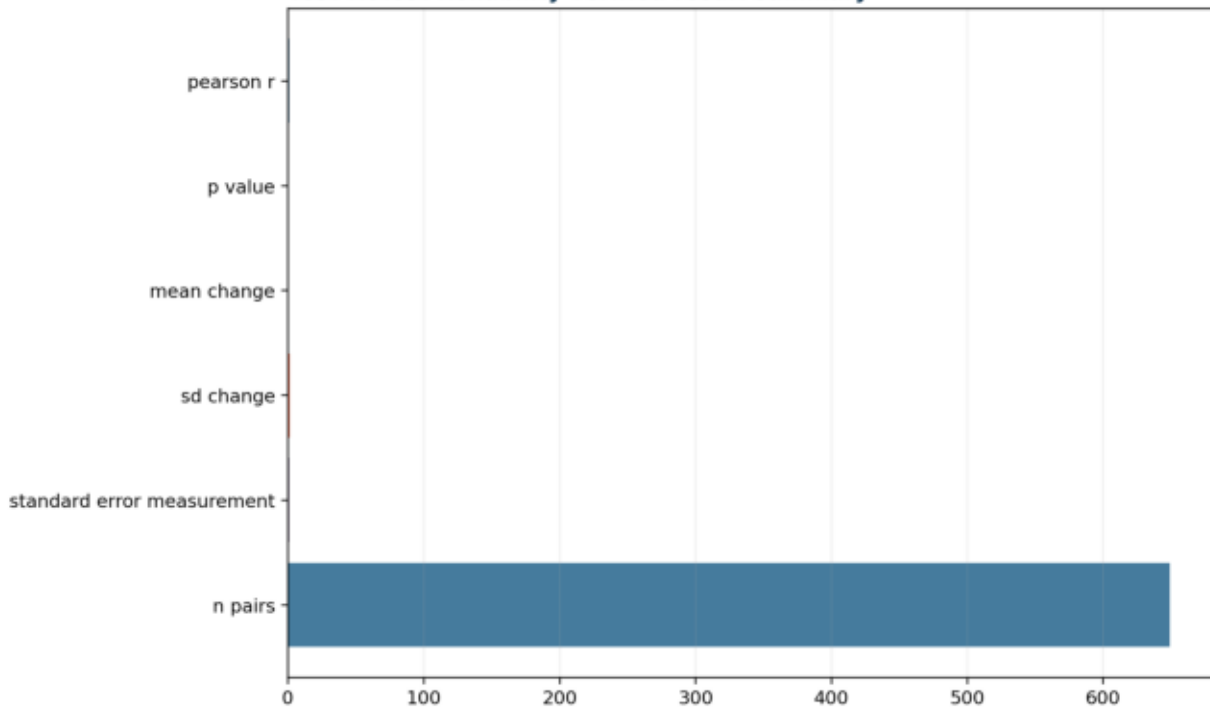
Change Score Summary



Stability Components



Test-Retest Reliability verified result summary



Primary Metrics

metric	value
pearson_r	0.864982
p_value	6.37379e-196
mean_change	0.171032
sd_change	1.47929
standard_error_measurement	1.04602
n_pairs	649

Paired Test Retest Scores

test_G1	retest_G2	change
0	11	11
9	11	2
12	13	1
14	14	0
11	13	2
12	12	0
13	12	-1
10	13	3
15	16	1
12	12	0
14	14	0
10	12	2
12	13	1
12	12	0
14	14	0
17	17	0
13	13	0
13	14	1
8	8	0
12	12	0
12	13	1
11	12	1
12	13	1
10	10	0
10	11	1
10	11	1

Change Score Summary

index	change
count	649
mean	0.171032
std	1.47929
min	-9
25%	-1
50%	0
75%	1
max	11

Stability Components

pearson_r	mean_change	SEM
0.864982	0.171032	1.04602