

1 When research does not synchronize:
2 Comprehensive analyses show no mother–child
3 physiological coupling - Supplementary material

4 Georgios Rousis^{1,2*}, Lisa Gistelinck^{1*}, Ward Deferm¹,
5 Rowena Van den Broeck¹, Bieke Bollen³, Sam Wass⁴,
6 Gunnar Naulaers³, Maarten De Vos², Bart Boets¹

7 ^{1*}Center for Developmental Psychiatry, KU Leuven, Herestraat 49,
8 Leuven, 3000, Belgium.

9 ²Department of Electrical Engineering, KU Leuven, Kasteelpark
10 Arenberg 10, Leuven, 3000, Belgium.

11 ³Department of Development and Regeneration, UZ Leuven,
12 Herestraat 49, Leuven, 3000, Belgium.

13 ⁴Developmental Psychology, University of East London, Water Lane,
14 London, E15 4L, United Kingdom.

15 *Corresponding author(s). E-mail(s): georgios.rousis@kuleuven.be;
16 lisa.gistelinck@kuleuven.be;

17 Contributing authors: ward.deferm@kuleuven.be;
18 rowena.vandenbroeck@kuleuven.be; bieke.bollen@uzleuven.be;
19 s.v.wass@uel.ac.uk; gunnar.naulaers@kuleuven.be;
20 maarten.devos@kuleuven.be; bart.boets@kuleuven.be;

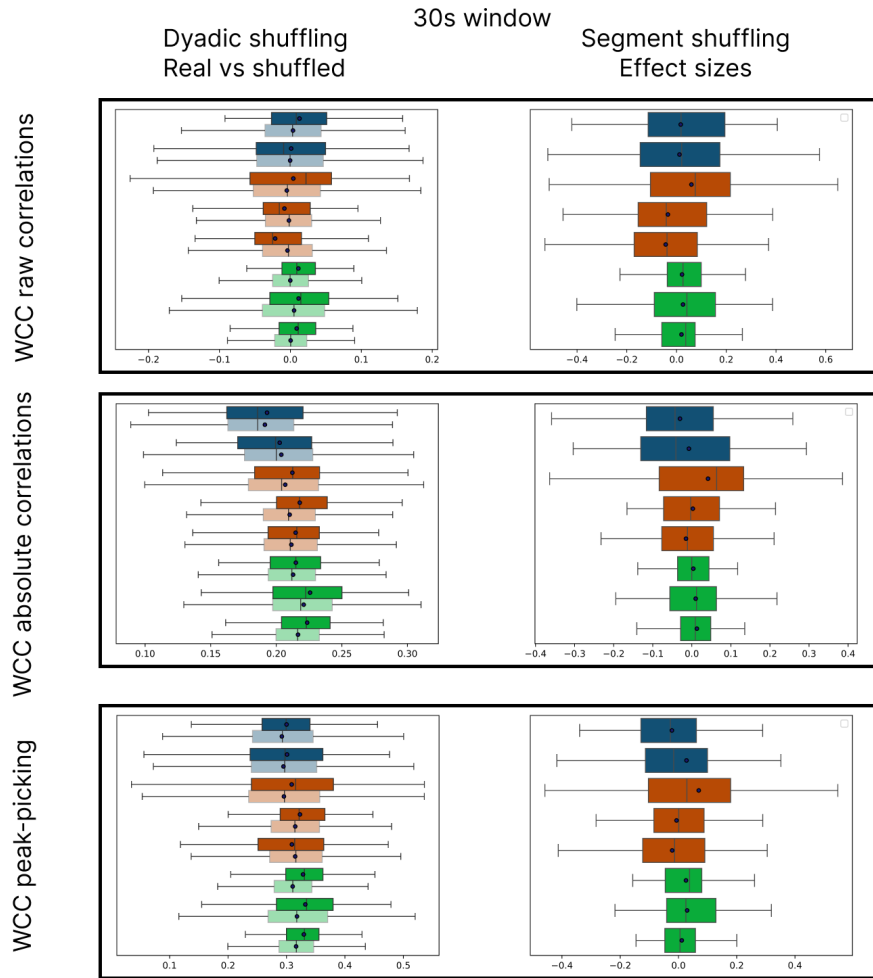
21 **Abstract**

22 Biobehavioral synchrony refers to covariation in physiological signals between
23 interacting partners and is widely hypothesized to support co-regulation, bond-
24 ing, and early socio-emotional development. This assumption is particularly
25 salient for children born preterm, who often experience reduced early physical
26 contact and prolonged neonatal separation, potentially shaping later care-
27 giver–child attunement. Here we present a large-scale, systematic investigation of
28 mother–child physiological synchrony in 102 dyads (70 preterm 5-year-olds and
29 32 matched full-term controls). Heart rate was recorded across eight interaction
30 paradigms spanning passive viewing to face-to-face play. We applied the major

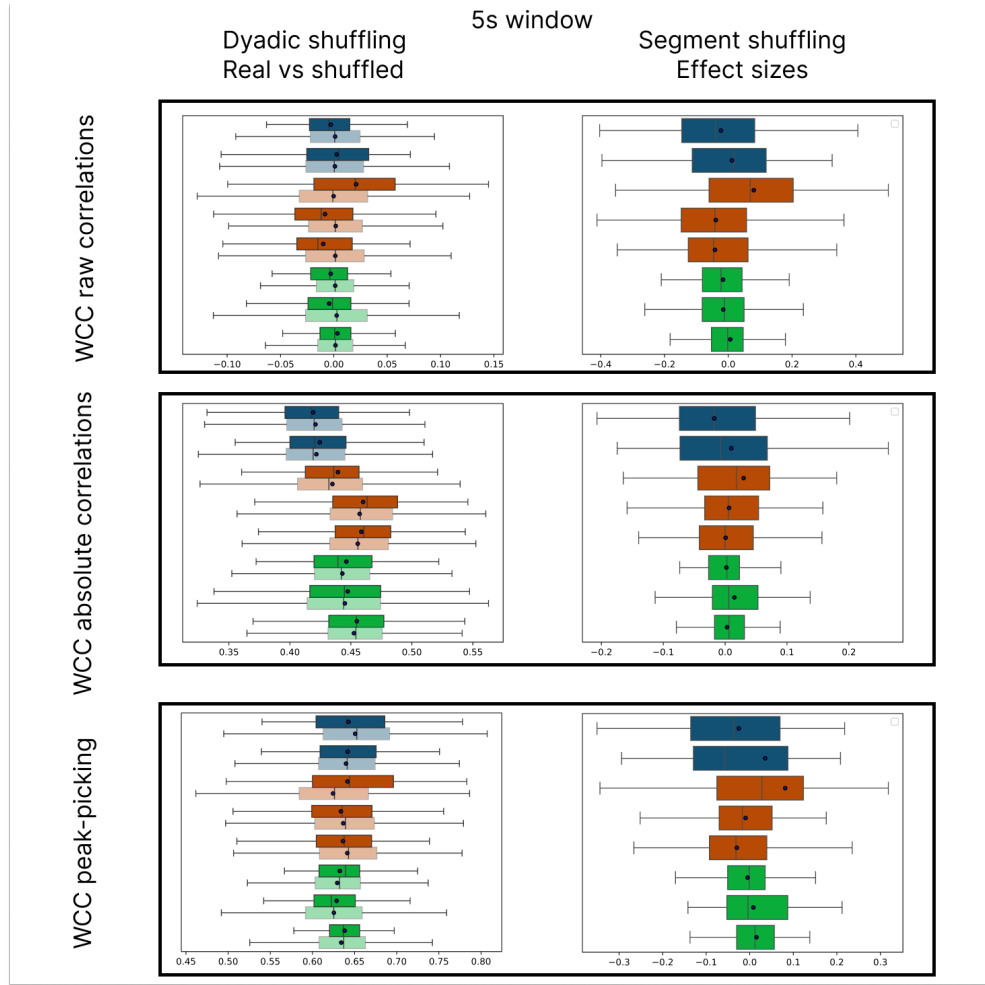
31 commonly used analytic approaches, combined with rigorous surrogate-data con-
32 trols and multiverse analyses, to assess the robustness of synchrony estimates.
33 Across all paradigms, metrics, and analytic variants, synchrony in real dyads
34 never exceeded that observed in pair-shuffled or segment-shuffled surrogate data.
35 Patterns were virtually identical in full-term and preterm groups, and no reliable
36 variability in synchrony remained that could be related to indices of prematurity.
37 Together, these findings challenge the view that autonomic synchrony is a robust
38 or ubiquitous feature of early social interaction and reveal substantial method-
39 ological fragility in current synchrony research. Our results underscore the need
40 for more rigorous analytical frameworks and raise critical questions about when
41 - and whether - physiological synchrony genuinely emerges in caregiver-child
42 relationships.

43 **Keywords:** Physiological synchrony, Prematurity, Mother-child interactions,
44 Multiverse analyses

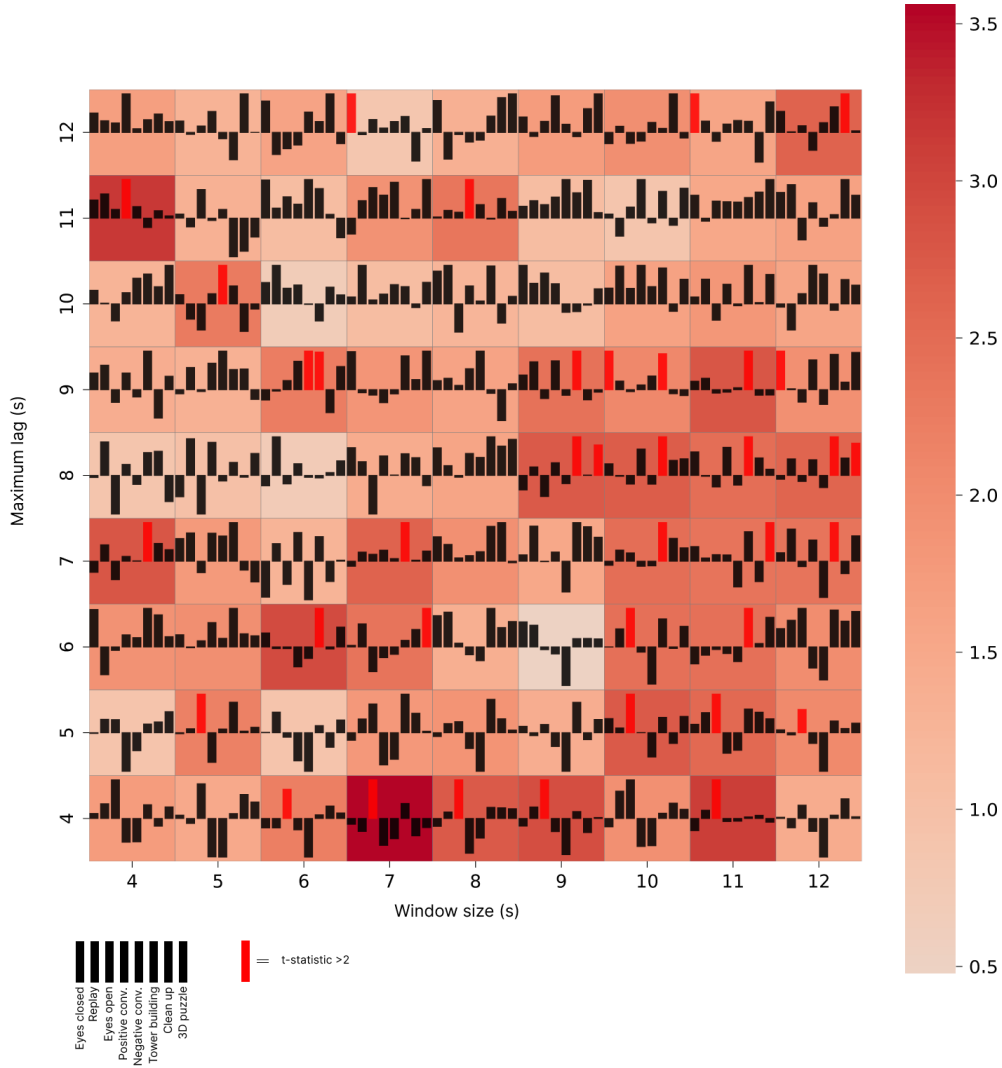
1 Supplementary figures



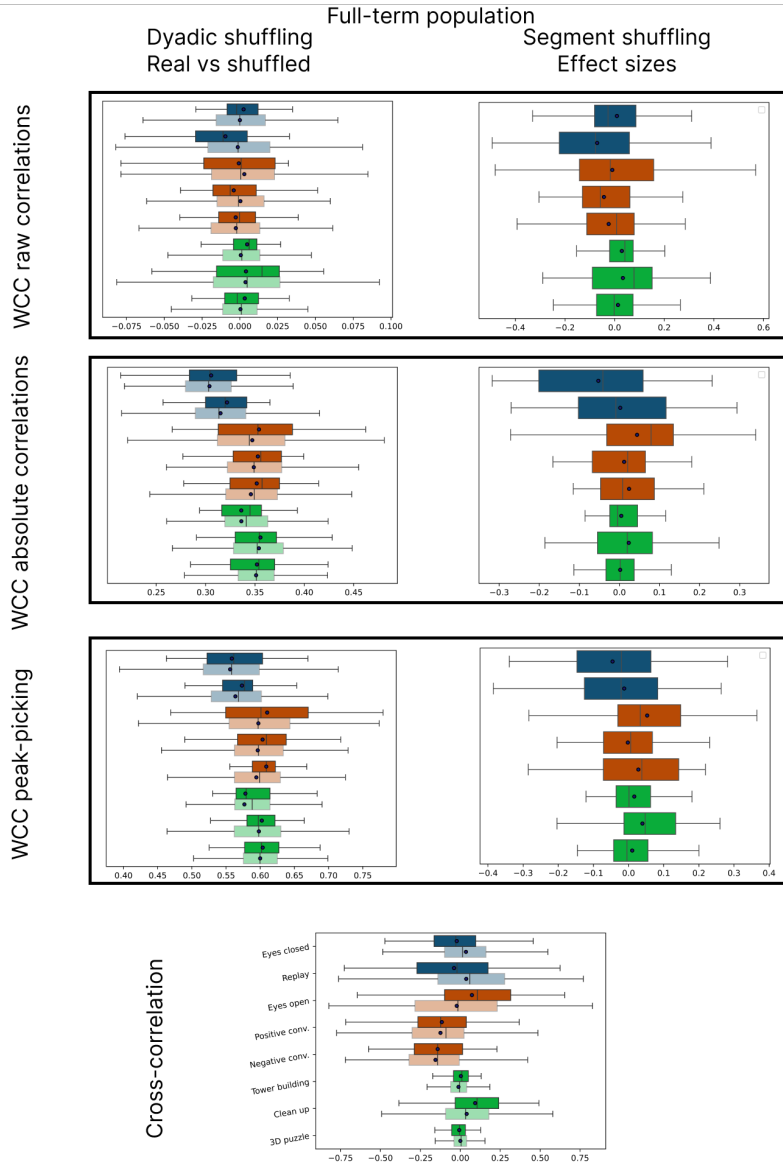
Suppl. Figure 1a Windowed time-lagged cross-correlations (WCC), computed using a 30 s sliding window with a 15 s time lag. Results are shown for the full sample across eight interaction conditions, grouped by interaction level (blue: no interaction; brown: box interaction; green: table interaction). The first column displays averaged raw correlations (A), absolute correlations (C), and peak-picking values (E) for real dyads (solid colors) and shuffled dyads (transparent color). The second column presents the corresponding effect sizes, quantifying differences between aligned and shuffled segments for each metric (B, D, F).



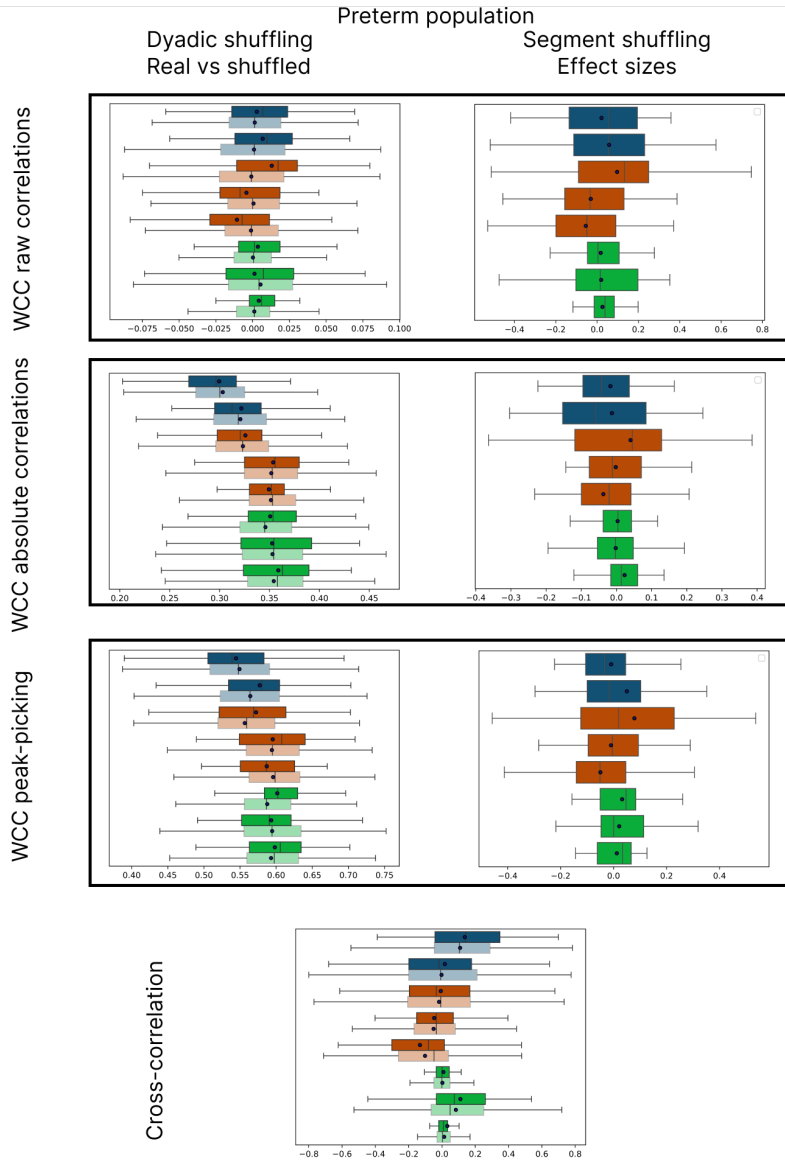
Suppl. Figure 1b Windowed time-lagged cross-correlations (WCC), computed using a 5s sliding window with a 2.5 s time lag. Results are shown for the full sample across eight interaction conditions, grouped by interaction level (blue: no interaction; brown: box interaction; green: table interaction). The first column displays averaged raw correlations (A), absolute correlations (C), and peak-picking values (E) for real dyads (solid colors) and shuffled dyads (transparent color). The second column presents the corresponding effect sizes, quantifying differences between aligned and shuffled segments for each metric (B, D, F).



Suppl. Figure 2 WCC synchrony across window sizes and time lags. Windowed cross-correlation (WCC) synchrony was evaluated across a parameter space comprising window sizes and maximum time lags ranging from 4 to 12 s. The figure displays two-sample t-statistics comparing real versus dyad-shuffled pairs across the full parameter space: (i) separately for each interaction condition (shown as black and red bar graphs within each cell), and (ii) averaged across the six real interaction conditions, excluding the eyes-closed and replay conditions (encoded by cell color). Each cell therefore contains bar graphs representing condition-specific t-statistics. Red bars indicate $t > 2$, and cell color reflects the maximum t-statistic observed across interaction conditions. Across the parameter space, synchrony patterns appeared largely random and unsystematic, and even the strongest effects remained well below the corrected significance threshold ($t > 4$).



Suppl. Figure 3a Cross-correlation and windowed time-lagged cross-correlation (WCC) for the full-term sample. Results are shown for the full-term sample across eight interaction conditions, grouped by interaction level (blue: no interaction; brown: box interaction; green: table interaction). WCCs were computed using a 10 s sliding window with a 5 s time lag. The first column displays averaged raw correlations (A), absolute correlations (C), and peak-picking values (E) for real dyads (solid colors) and shuffled dyads (transparent color). The second column presents the corresponding effect sizes, quantifying differences between aligned and shuffled segments for each metric (B, D, F). The bottom row illustrates cross-correlations computed over the full duration of each condition for real and shuffled dyads (G).



Suppl. Figure 3b Cross-correlation and windowed time-lagged cross-correlation (WCC) for the preterm sample. Results are shown for the preterm sample across eight interaction conditions, grouped by interaction level (blue: no interaction; brown: box interaction; green: table interaction). WCCs were computed using a 10 s sliding window with a 5 s time lag. The first column displays averaged raw correlations (A), absolute correlations (C), and peak-picking values (E) for real dyads (solid colors) and shuffled dyads (transparent color). The second column presents the corresponding effect sizes, quantifying differences between aligned and shuffled segments for each metric (B, D, F). The bottom row illustrates cross-correlations computed over the full duration of each condition for real and shuffled dyads (G).

2 Supplementary tables

Suppl. Table 1 Interacting vs non-interacting conditions - Absolute correlation comparisons

	Eyes closed		Replay	
	U-statistic	p-value	U-statistic	p-value
Eyes open	2328	2.164×10^{-6}	3482	0.082
Positive conv.	1470	1.108×10^{-13}	2277	6.6231×10^{-8}
Negative conv.	1431	1.428×10^{-12}	2158	2.5185×10^{-7}
Tower building	1428	1.3523×10^{-10}	2180	2.2106×10^{-5}
Clean up	1050	3.040×10^{-11}	1637	2.5687×10^{-6}
3D puzzle	1011	1.913×10^{-12}	1526	6.2056×10^{-8}

Suppl. Table 2 Interacting vs non-interacting conditions - Peak-picking value comparisons

	Eyes closed		Replay	
	U-statistic	p-value	U-statistic	p-value
Eyes open	3028	0.0069	3482	0.5655
Positive conv.	2332	6.406×10^{-7}	3891	0.0033
Negative conv.	2100	4.030×10^{-7}	3167	0.0071
Tower building	1880	7.547×10^{-7}	2998	0.0044
Clean up	1581	3.001×10^{-6}	2214	0.0089
3D puzzle	1531	2.360×10^{-7}	2186	0.0019

Suppl. Table 3 P-values of statistical surrogate tests - 30s window with 15s maximum lag

Condition	Dyadic shuffling ¹			Segment shuffling ²		
	Raw corr.	Abs. corr.	PP	Raw corr.	Abs. corr.	PP
Eyes closed	0.2212	0.7576	0.3806	0.8519	0.9311	0.9250
Replay	0.9190	0.8192	0.5598	0.9974	0.9508	0.9868
Eyes open	0.0974	0.2065	0.2069	0.9336	0.8919	0.9512
Positive conv.	0.2516	0.0147*	0.1271	0.8689	0.7943	0.9625
Negative conv.	0.0041*	0.2280	0.5309	0.7129	0.9123	0.9017
Tower building	0.0251*	0.4460	0.0011*	0.8060	0.9610	0.7576
Clean up	0.3458	0.4474	0.1218	0.8537	0.8281	0.7573
3D puzzle	0.0284*	0.0310*	0.0265*	0.8074	0.6731	0.7682

¹Comparison between real and shuffled data with Mann-Whitney U-tests.

²Comparison of effect sizes with 0 with one sample t-test.

Suppl. Table 4 P-values of statistical surrogate tests - 5s window with 2.5s maximum lag

Condition	Dyadic shuffling ¹			Segment shuffling ²		
	Raw corr.	Abs. corr.	PP	Raw corr.	Abs. corr.	PP
Eyes closed	0.2426	0.5531	0.1869	0.8984	0.8538	0.8663
Replay	0.6197	0.4738	0.9905	0.9583	0.9533	0.9476
Eyes open	0.0001*	0.4715	0.0236*	0.7241	0.8678	0.8708
Positive conv.	0.0116*	0.3805	0.7074	0.8175	0.9223	0.9372
Negative conv.	0.0164*	0.5043	0.3550	0.7793	0.9895	0.7733
Tower building	0.0750	0.6271	0.4239	0.8587	0.9633	0.9403
Clean up	0.2255	0.7690	0.7978	0.9041	0.8137	0.9357
3D puzzle	0.9385	0.6048	0.5416	0.9497	0.9384	0.8328

¹Comparison between real and shuffled data with Mann-Whitney U-tests.²Comparison of effect sizes with 0 with one sample t-test.**Suppl. Table 5** P-values of statistical surrogate tests - Full-term population

Condition	Dyadic shuffling ¹				Segment shuffling ²		
	CC	Raw corr.	Abs. corr.	PP	Raw corr.	Abs. corr.	PP
Eyes closed	0.1033	0.7657	0.7404	0.7783	0.9636	0.7224	0.8008
Replay	0.2106	0.118	0.2902	0.3879	0.7554	0.9892	0.9424
Eyes open	0.0570	0.6080	0.5527	0.5699	0.9682	0.7885	0.8017
Positive conv.	0.9488	0.2496	0.6118	0.4104	0.8125	0.9017	0.9839
Negative conv.	0.6692	0.9160	0.3679	0.2567	0.8829	0.8018	0.8427
Tower building	0.3470	0.2425	0.9581	0.8480	0.7862	0.9420	0.8267
Clean up	0.1519	0.6879	0.8608	0.6440	0.8679	0.8330	0.7860
3D puzzle	0.4350	0.7888	0.7797	0.8144	0.9253	0.9798	0.9221

¹Comparison between real and shuffled data with Mann-Whitney U-tests.²Comparison of effect sizes with 0 with one sample t-test.

Suppl. Table 6 P-values of statistical surrogate tests - Preterm population

Condition	Dyadic shuffling ¹				Segment shuffling ²		
	CC	Raw corr.	Abs. corr.	PP	Raw corr.	Abs. corr.	PP
Eyes closed	0.5157	0.4762	0.2299	0.5038	0.9247	0.9096	0.9637
Replay	0.9368	0.1408	0.8972	0.3218	0.8378	0.9567	0.9111
Eyes open	0.6426	0.0013*	0.9266	0.3651	0.7106	0.8937	0.8589
Positive conv.	0.7953	0.2715	0.7252	0.8519	0.8862	0.9822	0.9468
Negative conv.	0.2427	0.0554	0.6100	0.1383	0.7903	0.7274	0.7459
Tower building	0.6699	0.3278	0.3518	0.0321	0.9006	0.9559	0.7888
Clean up	0.5211	0.7918	0.8914	0.7008	0.9280	0.9840	0.8758
3D puzzle	0.5789	0.0970	0.3904	0.6667	0.7881	0.7046	0.8821

¹Comparison between real and shuffled data with Mann-Whitney U test.²Comparison of effect sizes with 0 with one sample t-test.**Suppl. Table 7** Full-term vs preterm - Effect size comparisons

Condition	Raw corr.		Abs corr.		PP	
	U-stat	p-value	U-stat	p-value	U-stat	p-value
Eyes closed	799	0.5644	773	0.4187	829	0.7579
Replay	636	0.0184*	1024	0.3404	869	0.7163
Eyes open	733	0.063	1072	0.3606	986	0.8344
Positive conv.	963	0.7260	1096	0.4908	1038	0.8163
Negative conv.	971	0.5181	1150	0.0278*	1169	0.0180*
Tower building	797	0.6567	758	0.9561	667	0.3987
Clean up	537	0.9487	607	0.4109	618	0.3352
3D puzzle	509	0.3100	460	0.1087	548	0.5446