

Figure S1: Root mean square of the single-participant SME values for the parent waves (a, b) and the difference waves (c, d). For the parent waves, the SME was computed only for mean amplitude (a) and peak amplitude (b), because latency values could not be validly measured from the parent waves in several cases. For the difference waves, the SME could be calculated for both the amplitude scores (c) and the latency scores (d).

Table S1: Mean (\pm standard error) of the average of SME values across participants, for each of the parent waves and for the difference waves for each of the seven ERP components.

Parent wave	P3b		N170		MMN		N400	
	Rare	Frequent	Faces	Cars	Deviants	Standards	Unrelated	Related
Mean amplitude	1.83 \pm 0.17	0.85 \pm 0.06	0.81 \pm 0.06	0.79 \pm 0.04	0.76 \pm 0.07	0.45 \pm 0.04	1.32 \pm 0.07	1.49 \pm 0.09
Peak amplitude	2.18 \pm 0.16	1.01 \pm 0.06	0.91 \pm 0.05	0.96 \pm 0.05	0.93 \pm 0.08	0.55 \pm 0.05	1.56 \pm 0.07	1.74 \pm 0.11
Parent wave	ERN		N2pc		LRP		-	
	Incorrect	Correct	Contralateral	Ipsilateral	Contralateral	Ipsilateral	-	-
Mean amplitude	1.88 \pm 0.12	0.58 \pm 0.02	0.47 \pm 0.03	0.47 \pm 0.03	0.59 \pm 0.03	0.59 \pm 0.03	-	-
Peak amplitude	2.59 \pm 0.21	0.70 \pm 0.02	0.55 \pm 0.03	0.55 \pm 0.03	0.67 \pm 0.03	0.68 \pm 0.03	-	-
Difference wave	P3b	N170	MMN	N400	ERN	N2pc	LRP	-
Mean amplitude	2.04 \pm 0.19	1.15 \pm 0.07	0.89 \pm 0.08	2.00 \pm 0.12	1.99 \pm 0.12	0.49 \pm 0.03	0.43 \pm 0.03	-
Peak amplitude	2.35 \pm 0.17	1.32 \pm 0.06	1.05 \pm 0.09	2.36 \pm 0.12	2.70 \pm 0.19	0.58 \pm 0.03	0.51 \pm 0.03	-
50% area latency	17.85 \pm 1.37	2.30 \pm 0.30	6.09 \pm 0.44	11.04 \pm 0.94	6.00 \pm 0.78	6.82 \pm 0.70	2.68 \pm 0.20	-
Peak latency	51.46 \pm 4.26	4.83 \pm 0.54	13.24 \pm 0.95	34.75 \pm 2.53	13.04 \pm 1.15	13.56 \pm 1.00	10.42 \pm 0.71	-

Table S2: Root mean square (RMS) of SME values across participants, for each of the parent waves and for the difference waves for each of the seven ERP components.

Parent wave	P3b		N170		MMN		N400	
	Rare	Frequent	Faces	Cars	Deviants	Standards	Unrelated	Related
Mean amplitude	2.13	0.93	0.90	0.83	0.88	0.53	1.38	1.6
Peak amplitude	2.40	1.08	0.97	1.01	1.05	0.63	1.62	1.86
Parent wave	ERN		N2pc		LRP		–	
	Incorrect	Correct	Contralateral	Ipsilateral	Contralateral	Ipsilateral	–	–
Mean amplitude	2.04	0.59	0.50	0.49	0.61	0.61	–	–
Peak amplitude	2.89	0.71	0.57	0.57	0.69	0.70	–	–
Difference wave	P3b	N170	MMN	N400	ERN	N2pc	LRP	–
Mean amplitude	2.35	1.23	1.02	2.13	2.13	0.52	0.46	–
Peak amplitude	2.57	1.38	1.19	2.48	2.96	0.61	0.54	–
50% area latency	19.79	2.97	6.68	12.50	7.72	8.06	2.95	–
Peak latency	57.93	5.89	14.50	38.18	14.87	14.90	11.32	–

Table S3: Paired t tests comparing SME or SD for the peak versus mean amplitude scoring methods for each ERP component, separately for each experimental condition (corrected for multiple comparisons across the family of 14 tests for each dependent variable). Note that the degrees of freedom were lower for the N2pc because Subject 7 was excluded from the N2pc analyses.

Parent wave	P3b		N170		MMN		N400		ERN		N2pc		LRP	
	t(39)	p	t(39)	p	t(39)	p	t(39)	p	t(39)	p	t(38)	p	t(39)	p
SME for condition 1	-7.13	<.001	-4.89	<.001	-12.00	<.001	-8.94	<.001	-6.41	<.001	-10.32	<.001	-12.19	<.001
SME for condition 2	-12.07	<.001	-9.35	<.001	-13.86	<.001	-7.97	<.001	-15.46	<.001	-11.08	<.001	-9.98	<.001
SD for condition 1	-3.74	0.001	-5.25	<.001	-10.60	<.001	-8.63	<.001	-7.49	<.001	-9.00	<.001	-7.66	<.001
SD for condition 2	-8.33	<.001	-8.33	<.001	-13.51	<.001	-5.70	<.001	13.60	<.001	-8.37	<.001	-7.48	<.001

Table S4: Paired t tests comparing the variance in the single-trial amplitudes between the two experimental conditions for each ERP component, separately for mean amplitude and peak amplitude (corrected for multiple comparisons across the family of tests for each scoring method). Note that the degrees of freedom were lower for the N2pc because Subject 7 was excluded from the N2pc analyses.

Parent wave	P3b		N170		MMN		N400		ERN		N2pc		LRP	
	t(39)	p	t(39)	p	t(39)	p	t(39)	p	t(39)	p	t(38)	p	t(39)	p
Variance for mean amplitude	-1.64	0.382	-0.14	0.904	-0.57	0.799	-2.75	0.063	0.12	0.904	0.63	0.799	-0.91	0.799
Variance for peak amplitude	-2.11	0.144	-1.27	0.371	-0.26	0.892	-1.90	0.152	2.76	0.063	0.14	0.892	-0.50	0.865

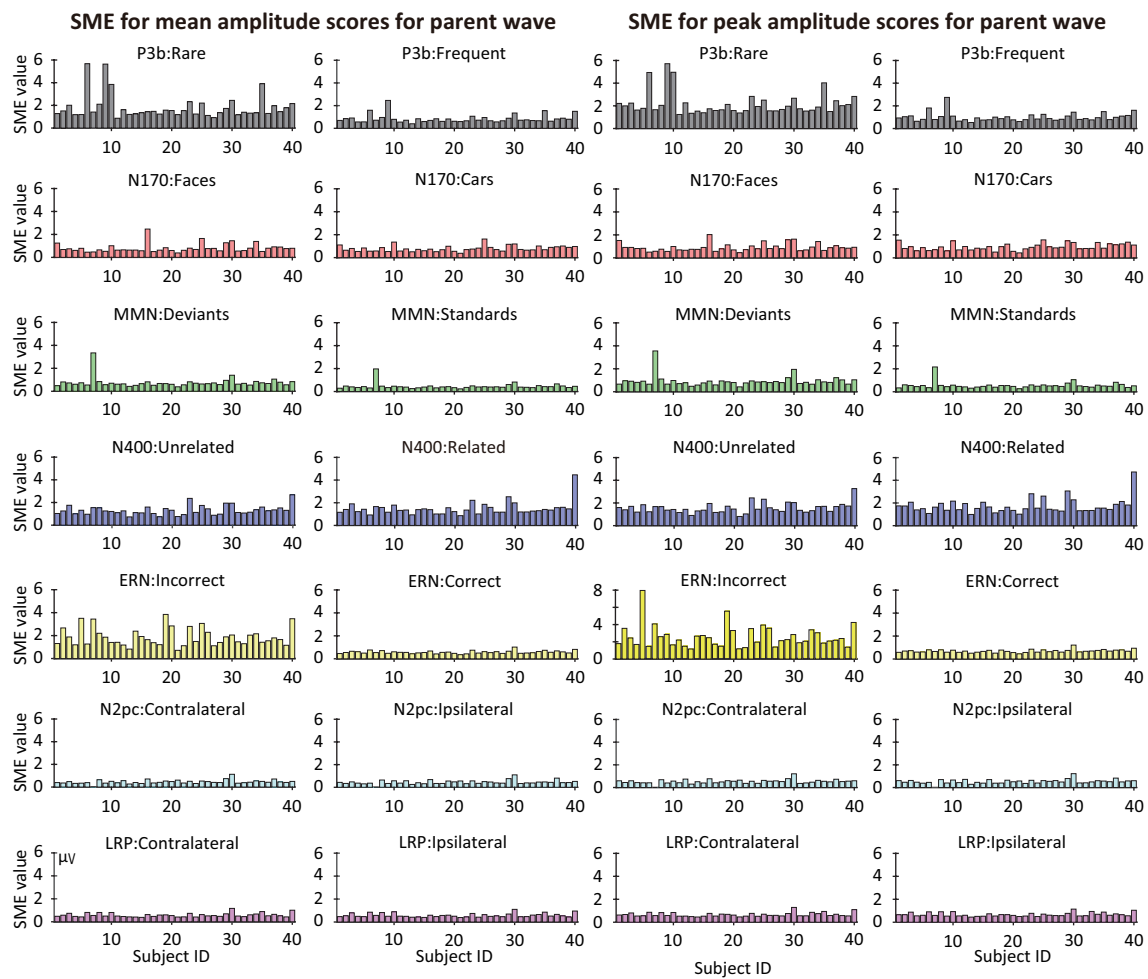


Figure S2: Single-participant SME values for mean amplitude and peak amplitude scores, quantified from the parent waves for each of the seven ERP components. Each bar represents one of the 40 participants. Values are provided only for mean amplitude and peak amplitude because latency values could not be validly measured from the parent waves in several cases.

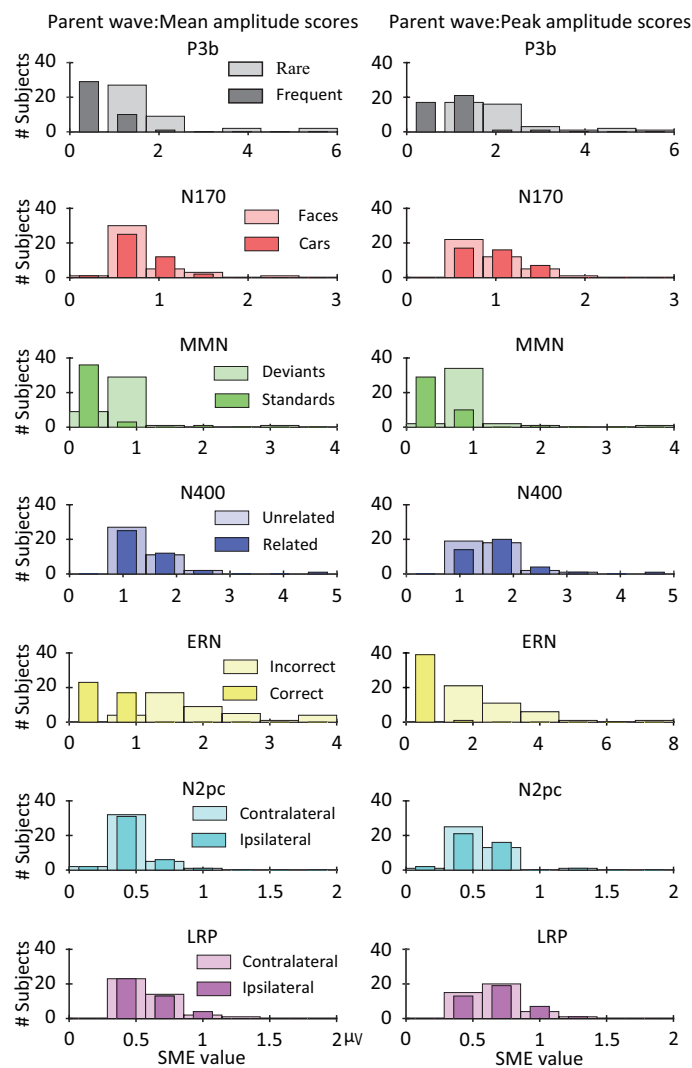


Figure S3: Histograms of single-participant SME values for mean amplitude and peak amplitude scores obtained from parent waves. For each component and each scoring method, the X axis was evenly divided into seven bins to reflect the different ranges of values for each plot. Values are provided only for mean amplitude and peak amplitude because latency values could not be validly measured from the parent waves in several cases.

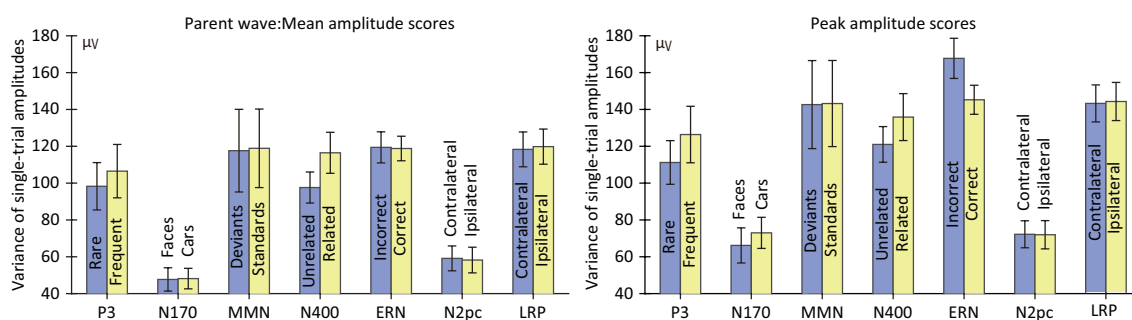


Figure S4: Analysis of the effects of variations in the number of trials on estimates of the variance in single-trial amplitude values. Panels c and d in Figure 4 of the main manuscript showed that the mean standard deviation (SD) across trials was numerically smaller in conditions with fewer trials than in conditions with more trials in several cases (for P3b mean and peak amplitude; for MMN mean and peak amplitude; for ERN mean amplitude). Although not statistically significant, these small differences in SD may reflect the fact that the equation typically used to estimate the SD is biased by the number of trials, underestimating (on average) the true SD by a progressively larger amount as the number of observations decreases (Gurland & Tripathi, 1971). As a result, the estimated SD for a condition with fewer trials will tend to be lower than the estimated SD for a condition with more trials, even if the true trial-to-trial variability is equivalent (all else being equal). The bias in the estimate of SD is a result of the square root operation used to convert the variance to the SD, and the equation for estimating the variance across trials is not biased by the number of trials. In the present figure, we show the variance rather than the SD across trials, averaged across participants for each condition. No significant differences between the rare and frequent conditions of the P3b, MMN, and ERP paradigms were present for the variance (see supplementary Table S4). Error bars show the standard error of the mean across participants.

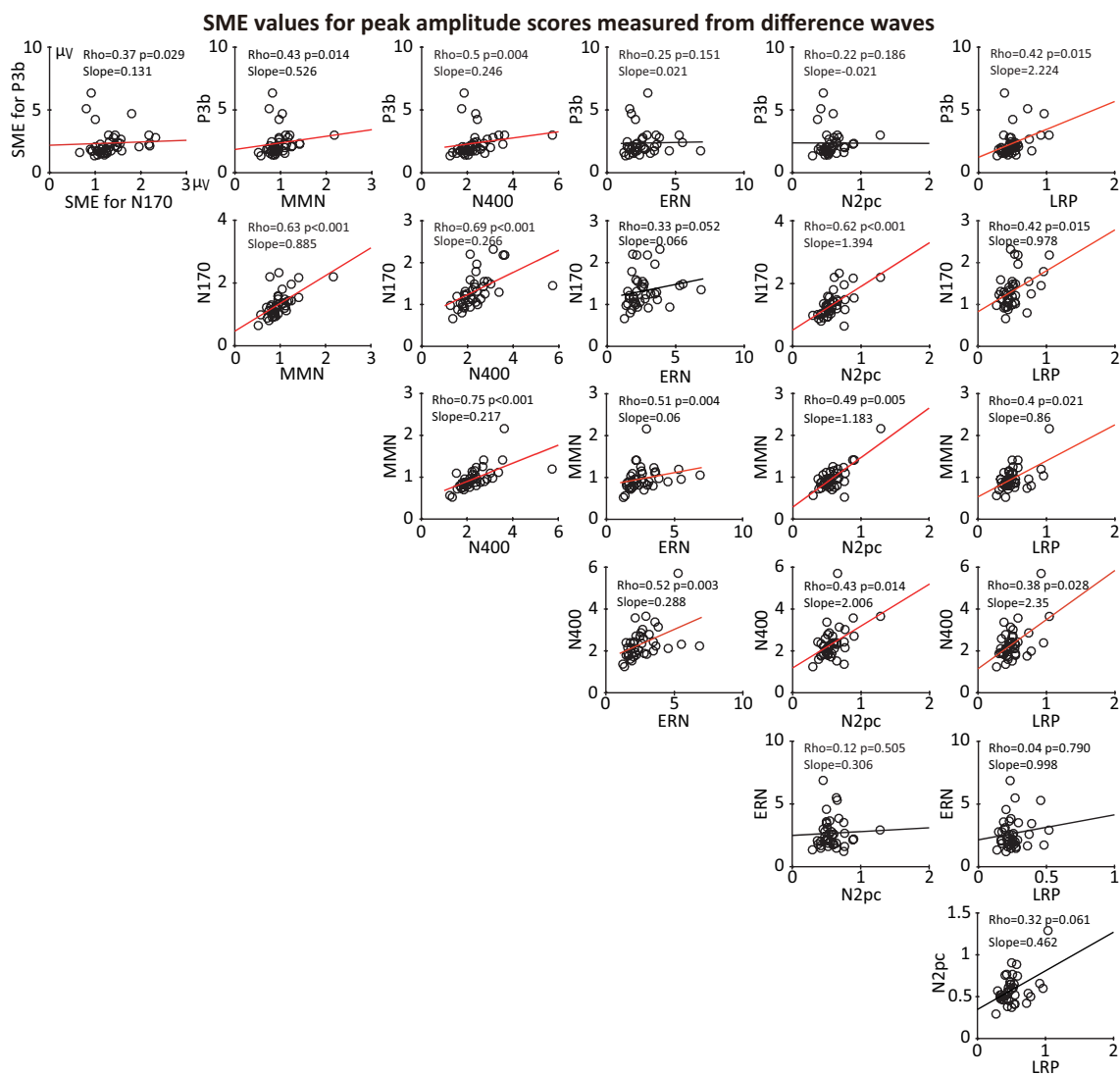


Figure S5: Scatterplots of the relationship between each pair of components for the SME values obtained for peak amplitude from the difference waves. Each circle represents a single participant. The p values were corrected for multiple comparisons across this entire family of tests. Note that Subject 7 was excluded from all correlations because the number of trials for one of conditions was zero for this participant in one condition of the N2pc paradigm.

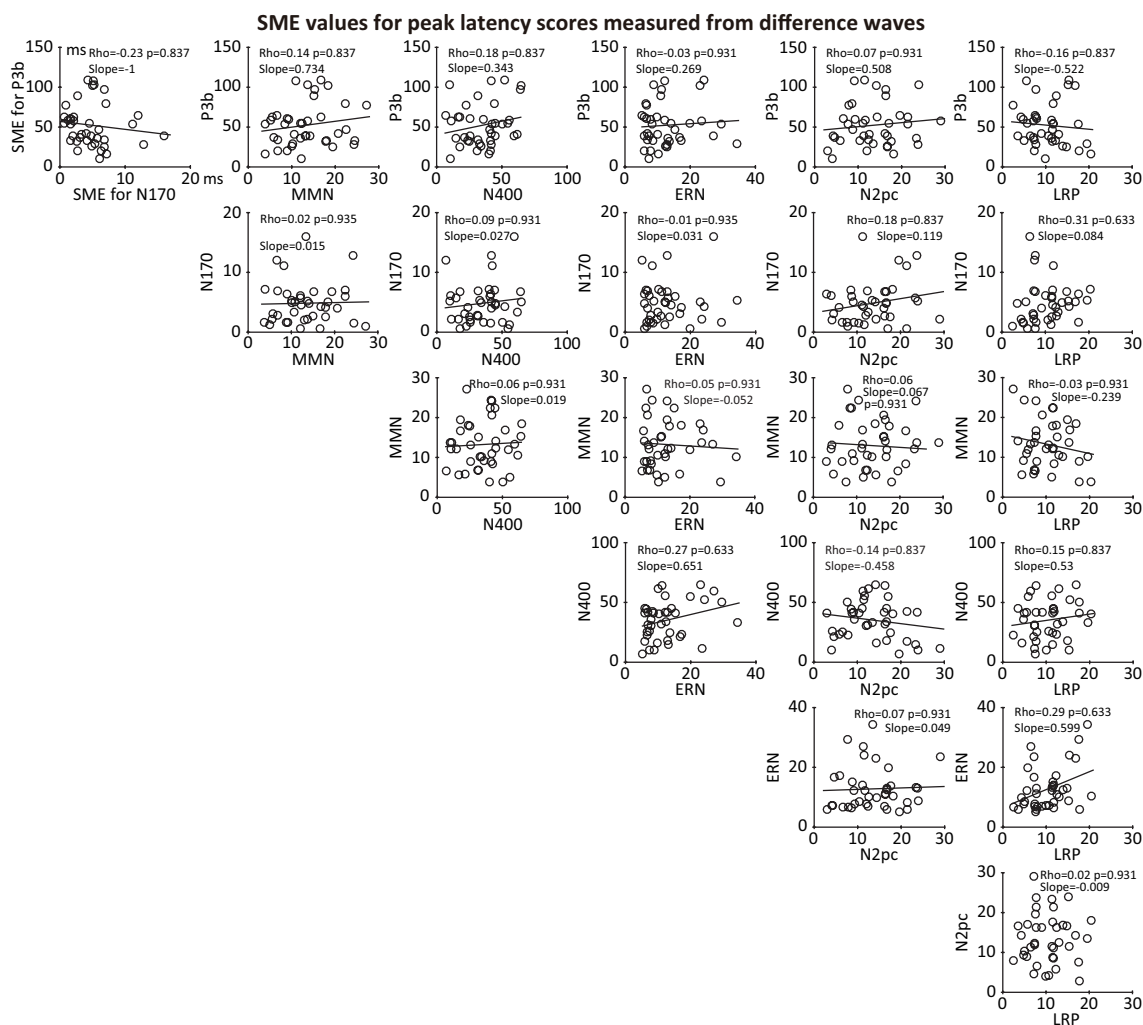


Figure S6: Scatterplots of the relationship between each pair of components for the SME values obtained for peak latency from the difference waves. Each circle represents a single participant. The p values were corrected for multiple comparisons across this entire family of tests. Note that Subject 7 was excluded from all correlations because the number of trials for one of conditions was zero for this participant in one condition of the N2pc paradigm.

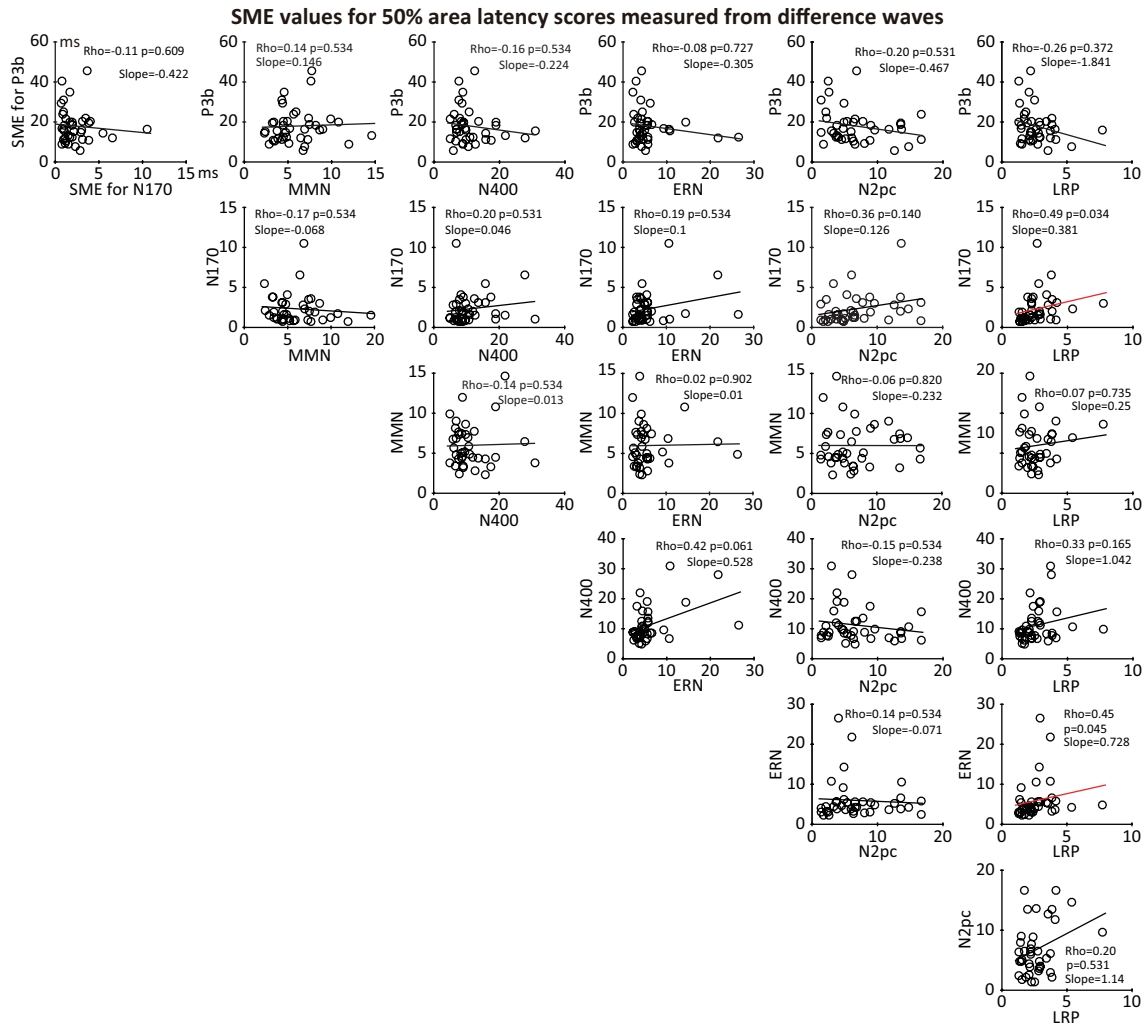


Figure S7: Scatterplots of the relationship between each pair of components for the SME values obtained for 50% area latency from the difference waves. Each circle represents a single participant. The p values were corrected for multiple comparisons across this entire family of tests. Note that Subject 7 was excluded from all correlations because the number of trials for one of conditions was zero for this participant in one condition of the N2pc paradigm.

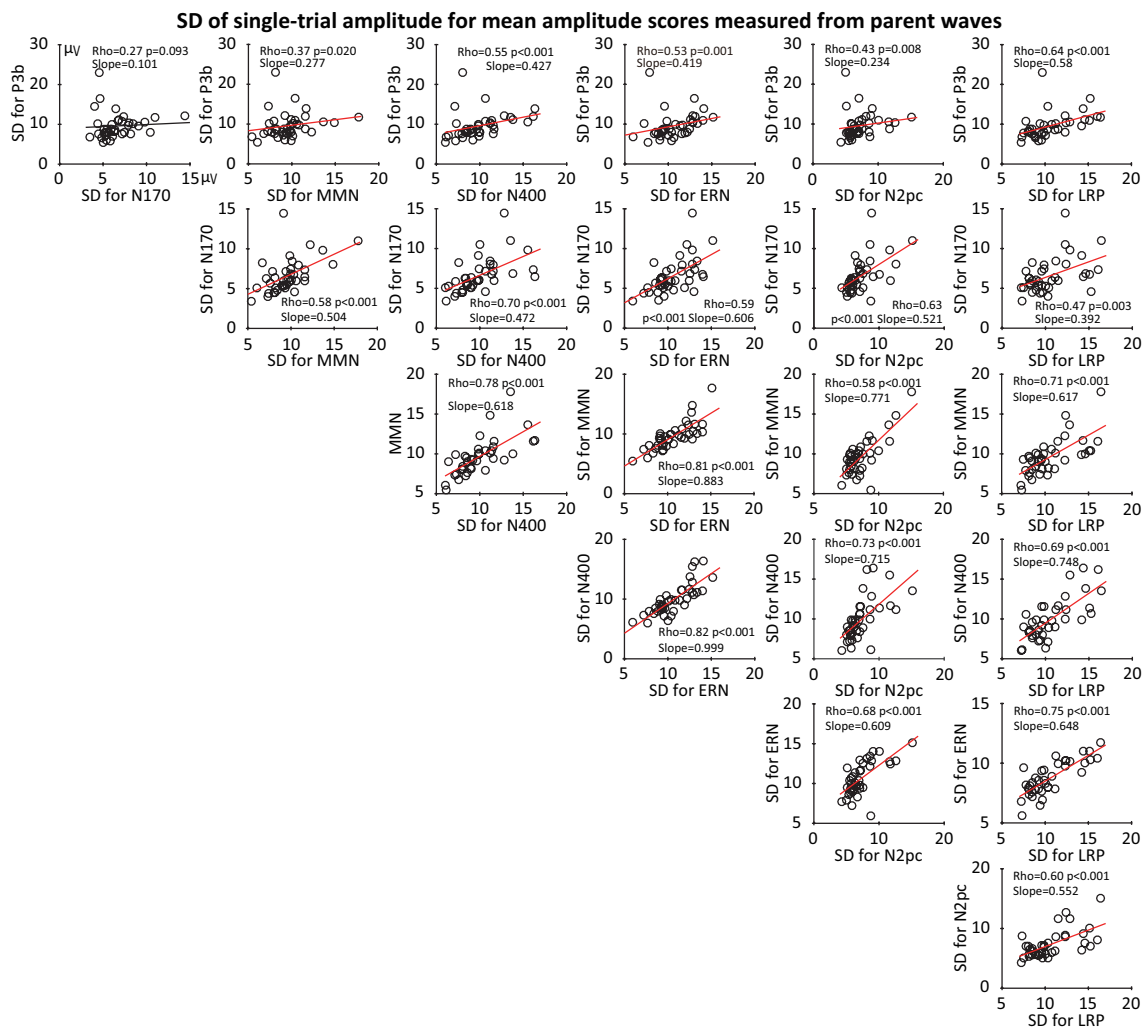


Figure S8: Scatterplots of the relationship between standard deviation (SD) values for mean amplitude for each pair of ERP components. The SD was calculated by measuring the mean amplitude from the single-trial epochs for a given condition and taking the SD of these values. We then averaged the SD across the two conditions used to define each component (because the SD values for the two conditions were strongly correlated, as shown in Figure S10). Each circle represents a single participant. The p values were corrected for multiple comparisons across this entire family of tests. Note that Subject 7 was excluded from all correlations because the number of trials for one of conditions was zero for this participant in one condition of the N2pc paradigm.

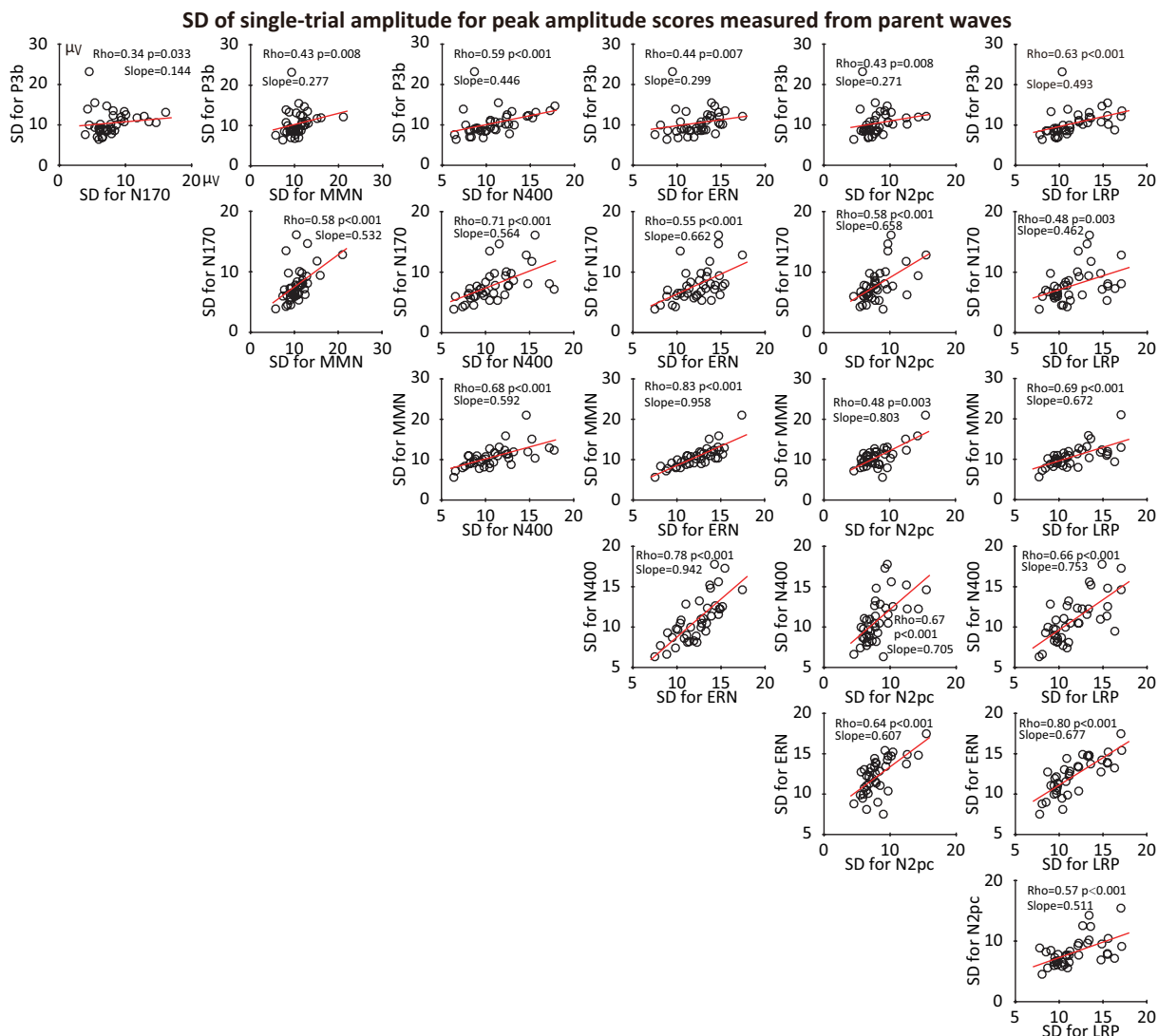


Figure S9: Scatterplots of the relationship between standard deviation (SD) values for peak amplitude for each pair of ERP components. The SD was calculated by measuring the mean amplitude from the single-trial epochs for a given condition and taking the SD of these values. We then averaged the SD across the two conditions used to define each component (because the SD values for the two conditions were strongly correlated, as shown in Figure S10). Each circle represents a single participant. The p values were corrected for multiple comparisons across this entire family of tests. Note that Subject 7 was excluded from all correlations because the number of trials for one of conditions was zero for this participant in one condition of the N2pc paradigm.

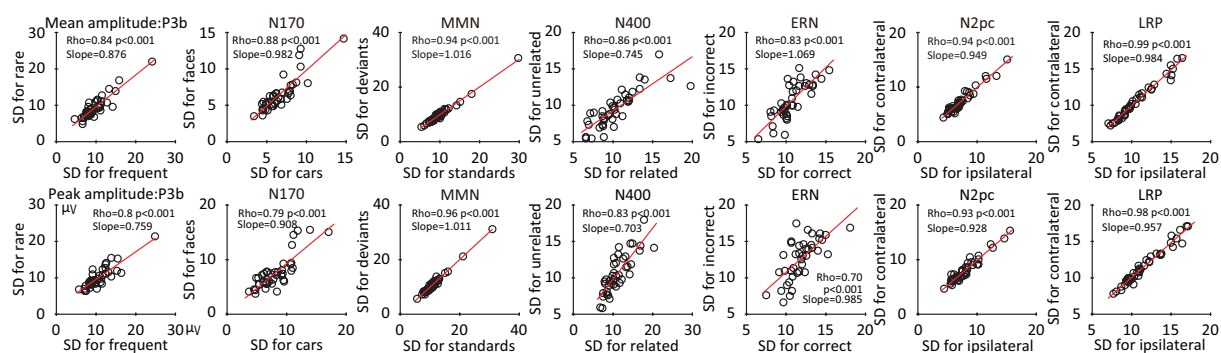


Figure S10: Scatterplots of the relationship between the SD values in the two conditions used to define each of the seven ERP components. The SD values were obtained from single-trial measurements of mean amplitude (top row) and peak amplitude (bottom row). Each circle represents a single participant. The p values were corrected for multiple comparisons across this entire family of tests.

766 **References**

- 767 Adrian, E. D., & Matthews, B. H. (1934). The berger rhythm: potential changes from the occipital lobes
768 in man. *Brain*, *57*, 355–385. doi:10.1093/brain/57.4.355.
- 769 Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J.
770 (2021). Power contours: Optimising sample size and precision in experimental psychology and human
771 neuroscience. *Psychological methods*, *26*, 295. doi:/10.1037/met0000337.
- 772 Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful
773 approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*,
774 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.
- 775 Boos, D. D. (2003). Introduction to the bootstrap world. *Statistical science*, *18*, 168–174. doi:10.1214/
776 ss/1063994971.
- 777 Boudewyn, M. A., Carter, C. S., Long, D. L., Traxler, M. J., Lesh, T. A., Mangun, G. R., & Swaab, T. Y.
778 (2017). Language context processing deficits in schizophrenia: The role of attentional engagement.
779 *Neuropsychologia*, *96*, 262–273. doi:10.1016/j.neuropsychologia.2017.01.024.
- 780 Brandmaier, A. M., Wenger, E., Bodammer, N. C., Kühn, S., Raz, N., & Lindenberger, U. (2018).
781 Assessing reliability in neuroimaging research through intra-class effect decomposition (ICED). *Elife*,
782 *7*, e35718. doi:10.7554/eLife.35718.
- 783 Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual
784 perception. *Journal of Neuroscience*, *29*, 7869–7876. doi:10.1523/JNEUROSCI.0113-09.2009.
- 785 Chernick, M. R. (2011). *Bootstrap methods: A guide for practitioners and researchers*. John Wiley &
786 Sons.
- 787 Clayson, P. E., Baldwin, S. A., & Larson, M. J. (2013). How does noise affect amplitude and latency
788 measurement of event-related potentials (ERPs)? a methodological critique and simulation study.
789 *Psychophysiology*, *50*, 174–186. doi:10.1111/psyp.12001.
- 790 Clayson, P. E., Brush, C. J., & Hajcak, G. (2021). Data quality and reliability metrics for event-related
791 potentials (erps): The utility of subject-level reliability. *International Journal of Psychophysiology*,
792 *165*, 121–136. doi:10.1016/j.ijpsycho.2021.04.004.
- 793 Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG
794 dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.
795 doi:10.1016/j.jneumeth.2003.10.009.
- 796 Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
- 797 Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter
798 in a nonsearch task. *Perception & psychophysics*, *16*, 143–149. doi:10.3758/BF03203267.

- 799 Gurland, J., & Tripathi, R. C. (1971). A simple approximation for unbiased estimation of the standard
800 deviation. *The American Statistician*, *25*, 30–32. doi:10.1080/00031305.1971.10477279.
- 801 Isbell, E., & Grammer, J. K. (2022). Event-related potentials data quality in young children: Standard-
802 ized measurement error of ERN and Pe. *Developmental Psychobiology*, *64*, e22245. doi:10.1002/dev.
803 22245.
- 804 Kappenman, E. S., Farrens, J. L., Zhang, W., Stewart, A. X., & Luck, S. J. (2021). ERP CORE: An
805 open resource for human event-related potential research. *NeuroImage*, *225*, 117465. doi:10.1016/j.
806 neuroimage.2020.117465.
- 807 Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and
808 statistical significance in ERP recordings. *Psychophysiology*, *47*, 888–904. doi:10.1111/j.1469-8986.
809 2010.01009.x.
- 810 Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-
811 related potentials. *Frontiers in Human Neuroscience*, *8*, 213. doi:10.3389/fnhum.2014.002.
- 812 Luck, S. J. (2005). An introduction to event related potentials and their neural origins. *An introduction*
813 *to the event related potential technique*, *11*.
- 814 Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.
- 815 Luck, S. J. (2022). *Applied Event-Related Potential Data Analysis*. LibreTexts. doi:10.18115/D5QG92.
- 816 Luck, S. J., Stewart, A. X., Simmons, A. M., & Rhemtulla, M. (2021). Standardized measurement error:
817 A universal metric of data quality for averaged event-related potentials. *Psychophysiology*, *58*, e13793.
818 doi:10.1111/psyp.13793.
- 819 Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see:
820 prestimulus α phase predicts visual awareness. *Journal of Neuroscience*, *29*, 2725–2732. doi:10.1523/
821 JNEUROSCI.3963-08.2009.
- 822 Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision
823 tasks. *Neural computation*, *20*, 873–922. doi:10.1162/neco.2008.12-06-420.
- 824 Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human
825 brain: Disentangling the contribution of low-level visual cues from face perception. *Vision research*,
826 *51*, 1297–1311.
- 827 Tamm, L., Narad, M. E., Antonini, T. N., O'Brien, K. M., Hawk, L. W., & Epstein, J. N. (2012). Reaction
828 time variability in ADHD: a review. *Neurotherapeutics*, *9*, 500–508. doi:10.1007/s13311-012-0138-5.
- 829 Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce ar-
830 tificial effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology*,
831 *52*, 997–1009. doi:10.1111/psyp.12437.

⁸³² Zhang, G., & Luck, S. J. (in preparation). An analysis of the factors that contribute to ERP data quality
⁸³³ and statistical power for seven common ERP components, .