Appendix

Measure	Time1 Mean (SD)	Time2 Mean (SD)	Time3 Mean (SD)	Exclusions *	ICC (2,k)	ICC (2,1)
Low reliability						
110% RMT Amp (mV)	0.44 (0.55)	0.36 (0.37)	0.34 (0.31)	32 (38)	0.37	0.17
110% RMT Area (mV ms)	2.55 (2.15)	1.99 (1.47)	1.61 (1.32)	35 (42)	0.32	0.13
125% AMT Area (mV ms)	13.76 (8.78)	10.06 (6.26)	7.69 (5.96)	38 (45)	0.27	0.11
150% AMT Area (mV ms)	25.89 (13.10)	19.51 (10.29)	17.26 (10.12)	44 (52)	0.39	0.17
175% AMT Area (mV ms)	36.15 (21.61)	27.02 (11.43)	22.64 (11.87)	48 (57)	0.21	0.08
125% AMT SPD (ms)	84.08 (35.59)	81.91 (28.29)	76.33 (31.86)	38 (45)	0.23	0.07
125% AMT SP ratio	0.16 (0.07)	0.12 (0.06)	0.10 (0.06)	38 (45)	0.40	0.18
150% AMT SPD (ms)	132.96 (45.96)	125.91 (36.19)	126.39 (39.95)	44 (52)	0.37	0.16
150% AMT SP ratio	0.19 (0.06)	0.15 (0.06)	0.14 (0.06)	44 (52)	0.41	0.19
175% AMT SP Ratio	0.21 (0.10)	0.17 (0.06)	0.15 (0.07)	48 (57)	0.16	0.06
SAF N34 Amplitude (mV)	1.34 (1.21)	1.28 (1.24)		45 (54)	0.48	0.32
SAI N22 Area (mV ms)	71.60 (38.71)	77.23 (33.27)		40 (48)	0.30	0.17
SAI N24 Area (mV ms)	78.50 (35.45)	85.22 (46.2)		40 (48)	0.02	0.01
SAF N34 Area (mV ms)	136.46 (78.22)	141.18 (73.27)		40 (48)	0.44	0.29
SAI N22 Ratio	0.79 (0.32)	0.89 (0.64)		46 (55)	0.49	0.32
SAF N32 Ratio	1.38 (0.61)	1.22 (0.54)		45 (54)	0.26	0.15
SAF N34 Ratio	1.37 (0.68)	1.20 (0.56)		45 (54)	0.15	0.08

Appendix Table 1: Denotes mean (standard deviation) values for dependent measures with low reliability (ICC(2,k) ≤ 0.5). For each measure, all time-points, as well as exclusions, ICC(2,k), and ICC(2,1) are shown. 'Exclusions' refers to participants who had to be excluded due to missing data (shown as count (% of total n)). * note that the Leiden site only recorded SEPs hence the large number of excluded RMT: resting motor threshold, AMT: active motor threshold, SPD: silent period duration, SAI: short-latency afferent inhibition, SAF: short afferent facilitation. 22, 24, 32, 34 represent interstimulus intervals between the nerve stimulation and TMS.

Low Reliability							
110 RMT Area (mV ms)	0.43	6, 102	<0.001	1.90 (0.92)	1.50 (0.86)	2.33 (1.06)	N/A
125 AMT Area (mV ms)	0.41	6, 96	<0.001	11.92 (4.30)	10.44 (5.48)	8.46 (3.46)	N/A
150 AMT Area (mV ms)	0.52	6, 86	<0.001	22.63 (7.29)	19.55 (7.35)	17.61 (8.01)	N/A
175 AMT Area (mV ms)	0.39	6, 80	0.01	30.67 (13.41)	27.15 (7.55)	25.89 (10.24)	N/A
125 AMT SPD (ms)	0.39	6, 96	0.002	86.04 (17.29)	84.27 (18.09)	74.35 (17.26)	N/A
150 AMT SPD (ms)	0.59	6, 86	<0.001	132.97 (28.33)	131.42 (23.13)	114.17 (29.92)	N/A
125 AMT Ratio	0.36	6, 96	0.003	0.13 (0.04)	0.12 (0.05)	0.12 (0.04)	N/A
150 AMT Ratio	0.53	6, 86	<0.001	0.17 (0.04)	0.15 (0.04)	0.15 (0.04)	N/A
SAI 24 Area (mV ms)	0.20	4, 96	0.04	0.96 (0.57)	0.87 (0.60)	0.66 (0.74)	N/A
SAF 34 Ratio	0.23	4, 86	0.03	1.40 (0.43)	1.18 (0.33)	1.09 (0.47)	N/A

Appendix Table 2. This table shows significant effects study site in a MANOVA for electrophysiological measures with low reliability. Mean (SD) are shown for each study site averaging across time points. Only participants with data from all three time-points were included in this analysis. Leiden data was only available for EEG measures. RMT: resting motor threshold; AMT: active motor threshold; SPD: silent period duration; SAI: short-latency afferent inhibition; SAF: short-latency afferent facilitation. 22, 24, 32, 34 denote interstimulus intervals between the nerve stimulation and TMS.

Measure	V	df	p	London	Paris	Vancouver	Leiden
High Reliability							
RMT (%MSO)	0.20	6, 108	0.07	45.73 (10.55)	42.60 (8.32)	49.31 (8.54)	N/A
AMT (%MSO)	0.12	6, 110	0.34	36.32 (7.42)	33.48 (6.02)	37.52 (6.87)	N/A
Active Latency (ms)	0.15	6, 94	0.26	20.07 (1.77)	21.00 (1.82)	21.38 (1.34)	N/A
150 RMT Amp (mV)	0.09	6, 92	0.67	2.50 (1.83)	2.42 (1.89)	1.79 (1.63)	N/A
N20 Latency (ms)	0.17	9, 189	0.27	19.53 (1.26)	20.13 (1.27)	19.86 (0.79)	20.20 (1.19)
LLR1 Latency	0.20	6, 68	0.29	37.05 (3.26)	39.12 (3.66)	38.15 (1.83)	N/A
LLR2 Latency	0.25	6, 68	0.15	47.86 (3.51)	50.41 (3.65)	50.13 (2.39)	N/A
Moderate Reliability							
SEP Amp ST (μV)	0.23	9, 165	0.15	0.74 (0.37)	0.42 (0.47)	0.76 (0.38)	0.93 (0.72)
LLR2 Amp (mV)	0.15	6, 70	0.49	0.10 (0.05)	0.08 (0.03)	0.07 (0.07)	N/A
CRT (ms)	0.19	6, 54	0.47	7.81 (1.85)	8.10 (1.85)	9.56 (0.95)	N/A
130 RMT Amp (mV)	0.10	6, 98	0.53	1.34 (1.04)	1.44 (1.11)	1.18 (0.99)	N/A
175 AMT Amp (mV)	0.20	6, 82	0.19	4.44 (1.32)	4.02 (2.05)	4.96 (3.19)	N/A
150 RMT Area (mV ms)	0.20	6, 88	0.14	10.87 (7.35)	12.36 (7.54)	6.68 (5.23)	N/A
SAI22 Amp (mV)	0.02	4, 84	0.94	79.02 (28.71)	70.06 (24.94)	68.28 (24.53)	N/A
SAI24 Amp (mV)	0.05	4, 86	0.71	93.79 (32.42)	74.77 (18.94)	70.64 (28.47)	N/A
SAF32 Amp (mV)	0.14	4, 86	0.18	163.78 (60.94)	124.04 (44.22)	114.20 (69.06)	N/A
SAF34 Amp (mV)	0.13	4, 86	0.22	160.58 (59.81)	136.70 (70.98)	112.11 (62.38)	N/A
Low Reliability							
110 RMT Amp (mV)	0.11	6, 106	0.42	0.31 (0.24)	0.43 (0.33)	0.41 (0.29)	N/A
175 AMT Ratio	0.29	6, 80	0.15	0.19 (0.07)	0.16 (0.04)	0.18 (0.04)	N/A
SAI22 Area (mV ms)	0.07	4, 96	0.43	0.88 (0.55)	0.75 (0.54)	0.77 (0.85)	N/A
SAF32 Area (mV ms)	0.15	4, 96	0.12	1.46 (0.66)	1.49 (1.23)	1.00 (1.08)	N/A
SAF34 Area (mV ms)	0.09	4, 96	0.33	1.49 (0.74)	1.38 (1.09)	0.88 (0.96)	N/A
SAI22 Ratio	0.07	4, 84	0.55	0.76 (0.35)	0.82 (0.50)	0.99 (0.47)	N/A
SAI24 Ratio	0.12	4, 86	0.25	0.90 (0.36)	0.83 (0.34)	0.83 (0.48)	N/A
SAI32 Ratio	0.14	4, 86	0.18	1.40 (0.45)	1.25 (0.40)	1.23 (0.40)	N/A

Appendix Table 3. This table shows non-significant study site effects from MANOVA results. For latency variables, the MANOVA controlled for the arm length of each participant. Mean (SD) are shown for each study site averaging across time points. Only participants with data from all three time-points were included in this analysis. Leiden data was only available for EEG measures. SEP: somatosensory evoked potential; ST: sensory threshold; LLR: long-latency reflex; CRT: cortical relay time; RMT: resting motor threshold; AMT: active motor threshold; SAI: short-latency afferent inhibition; SAF: short-latency afferent facilitation. 22, 24, 32, 34 denote interstimulus intervals between the nerve stimulation and TMS.

Post-Hoc ANOVAs for Significant Between-Site Effects

High Reliability

For **SEP Amp MT**, at *Time 1*, Leiden was significantly different from London and Paris (ps < 0.006), but not Vancouver (p = 0.24). Neither London, Paris, nor Vancouver were statistically different from each other. At *Time 2*, Leiden was again different from London and Paris (ps < 0.005), but not Vancouver (p = 0.15). Neither London, Paris, nor Vancouver were statistically different from each other. At *Time 3*, Leiden was significantly different from London, Paris, and Vancouver (ps < 0.04), but these three groups were not statistically different from each other.

For **SEP Amp 150**, at *Time 1*, Leiden was significantly different from London and Paris (ps < 0.001), but not from Vancouver (p = 0.94). At *Time 2*, Leiden was significantly different from London and Paris (ps < 0.011), but none of the other sites were statistically different from each other. At *Time 3*, Leiden was again statistically different from London and Paris (ps < 0.04), London was statistically different from Paris (p 0.03), and Vancouver was not statistically different from any of the other study sites.

For **Rest Latency**, at *Time 1*, there was a significant difference between Paris and Vancouver (p = 0.02), but there were no other significant differences between study sites (ps > 0.11). At *Time 2*, there was a significant difference between London and Paris (p = 0.005), but there were no other significant differences between study sites (ps > 0.19). For *Time 3*, there was a significant difference between London and Paris (p = 0.001), but there were no other significant differences between study sites (ps > 0.09).

Moderate Reliability

For **125 AMT Amp**, at *Time 1*, London was significantly different from Paris and Vancouver (ps < 0.001), which were not different from each other (p = 0.59). At *Time 2*, there were no significant differences between sites (ps > 0.26). At *Time 3*, there were no significant differences between study sites (ps > 0.05).

For **150 AMT Amp**, at *Time 1*, London was significantly different from Paris and Vancouver (ps = 0.001), which were not significantly different from each other (p = 0.63). At *Time 2*, there were no significant differences between study sites (ps > 0.53). At *Time 3*, there were no significant differences between study sites (ps > 0.71).

For **175 AMT SPD**, at *Time 1*, Vancouver was significantly different from London (p < 0.001) and Paris (p < 0.001), which were not different from each other (p = 0.72). At *Time 2*, there were no significant differences between study sites (ps > 0.44). At *Time 3*, there was a significant difference between London and Paris (p = 0.04) but no other significant differences between study sites (ps > 0.16).

Low Reliability

For **110 RMT Area**, at *Time 1*, there were no significant differences between study sites (ps >0.13). At *Time 2*, London was significantly different from Paris (p = 0.02) and Vancouver (p = 0.02), which were significantly different from each other (p < 0.001). At *Time 3*, Vancouver was different from London (p = 0.01) and Paris (p = 0.03), which were no different from each other (p = 0.73).

For **130 RMT Area**, at *Time 1*, Paris was significantly different from London (p = 0.4) and Vancouver (p = 0.003), which were not different from each other (p = 0.24). At *Time 2*, there were no significant differences between sites (ps > 0.43). At *Time 3*, there were no significant differences between groups (ps > 0.31).

For **AMT 125 Area**, at *Time 1*, Vancouver was significantly different from London and Paris (ps < 0.001), but London and Paris were not different from each other (p = 0.86). At *Time 2*, there were no significant differences between study sites (ps > 0.06). At *Time 3*, there were also no significant differences between study sites (ps > 0.34).

For **AMT 150 Area**, at *Time 1*, there was a significant difference between London and Vancouver (p < 0.001), but no significant difference between London and Paris (p = 0.39). At *Time 2*, there were no statistically significant differences between groups (ps > 0.20). At *Time 3*, there was a statistically significant difference between Vancouver and London (p = 0.04) and Paris (p = 0.03), but no statistically significant difference between London and Paris (p = 0.71).

For **AMT 175 Area**, at *Time 1*, there was a significant difference between Vancouver and London (p = 0.03), but not London and Paris (p = 0.68). At *Time 2*, there were no statistically significant differences between sites (ps > 0.09). At *Time 3*, Vancouver was significantly different from London (p = 0.008) and Paris (p = 0.009), but London and Paris were not different from each other (p = 0.94).

For **125 AMT SPD**, at *Time 1*, Vancouver was significantly different from London (p < 0.001) and Paris (p = 0.03), which were also different from each other (p = 0.01). At *Time 2*, there was no statistically significant differences between groups (ps > 0.63). At *Time 3*, London was significantly different from Vancouver (p = 0.04) and Paris (0.03), which were not significantly different from each other.

For **150 AMT SPD**, at *Time 1*, Vancouver was significantly different from London (p < 0.001) and Paris (p = 0.002), which were different from each other (p = 0.004). At *Time 2*, there were no statistically significant differences between study sites (ps > 0.74). At *Time 3*, there we no statistically significant differences between study sites (ps > 0.05).

For **125 AMT Ratio**, at *Time 1*, Vancouver was significantly different from Paris (p = 0.001), but no other differences were significant (p > 0.05). At *Time 2*, London and Paris were significantly different (p = 0.02), but no other differences were significant (ps > 0.09). At *Time 3*, there were no statistically significant differences between groups (ps > 0.05).

For **150 AMT Ratio**, at *Time 1*, Vancouver was significantly different from London (p = 0.03) and Paris (p = 0.005), which were not different from each other (p = 0.30). At *Time 2*, London and Paris were significantly different from each other (p = 0.02), but no other differences were significant (ps > 0.06). At *Time 3*, Vancouver was significantly different from Paris (p = 0.01), but no other differences were statistically significant (ps > 0.11).

For **SAI 24 Area**, at *Time 1*, there were no significant differences between groups (ps > 0.88). At *Time 2*, London was significantly different from Paris (p = 0.005) and Vancouver (p = 0.04), which were not different from each other (p = 0.66).

For **SAF 34 Ratio**, at *Time 1*, Vancouver was not significantly different from London (p = 0.41) or Paris (0.24), although those two sites were different from each other (p = 0.03). At *Time 2*, Vancouver was significantly different from both London (p = 0.04) and Paris (p = 0.04), but these sites were not different from each other (p = 0.91).