Supplementary Material

Supplementary material to Heggli, O.A., Konvalinka, I., Kringelbach, M.L., Vuust, P. *A Metastable Attractor Model of Self-Other Integration (MEAMSO) in Rhythmic Synchronization.* Philosophical Transactions of the Royal Society B (2021).

Appendix 1: Detailed description of the calculations in the MEAMSO

The main output of the MEAMSO is an integer, $brainstate_i$, indicating the current configuration (integration or segregation) at event i. This integer is calculated as the rounded value of a Bayesian average, dependent on a constant C, an updating prior p_i, and the current and previous instantaneous outcome of the model selection function of the MEAMSO, here denoted $state_i$,

$$brainstate_i = \left\lfloor \frac{C \times p_i + \sum_{i=1}^{N} state_i}{C + N} \right\rceil \tag{1}$$

with N determining how many previous events should be considered. The prior p is updated for each round of calculations, by adding to the previous prior the absolute difference between p_{i-1} and $brainstate_{i-1}$, divided by a scaling variable s as shown,

$$p_i = p_{i-1} + \frac{|p_{i-1} - brainstate_{i-1}|}{s} \tag{2}$$

The model selection function indicates whether the current input to MEAMSO is considered to be consistent with integrating (one model) or segregating (two models) between self and other at event i, here denoted as $state_i$. This function is an argmax, selecting the vector position of highest value in an array (σ_1, σ_2) where σ_1 is the scaled evidence for integrating, and σ_2 is the scaled evidence for segregating as shown in the following equation: $state_i = argmax(\sigma_1, \sigma_2)$. Here, σ is scaled using a softmax function $\sigma(fit)_i = \frac{e^{fit_i}}{\sum_{j=1}^{2} e^{fit_j}}$, with fit being a vector containing the evidence for the two competing models $fit = (evid_{one}, evid_{two})$. For

$evid_{one}$ we collect and add any positive values from the individual comparisons calculated in the MEAMSO, as shown,

$$evid_{one} = K \sum_{i=1}^{4} comp_i, \;\; for\ comp_i > 0 \tag{3}$$

For $evid_{two}$ we collect and add any negative values from the individual comparisons with the exception that comparison values are summarized by their absolute value, as shown,

$$evid_{two} = K \sum_{i=1}^{4} |comp_i|, \;\; for\ comp_i < 0 \tag{4}$$

Here $comp = (comp_{time}, comp_{aud}, comp_{nback}, comp_{corr})$, and K is a scaling variable. These individual comparisons are described in the main paper, but a short summary follows here.

$comp_{time}$ is the more-or-less instantaneous comparison between a self-initiated finger tap and the potentially causally linked auditory feedback. This comparison is performed by calculating the asynchrony, the time difference between the two events. Here, we apply a reversed logistic function for asynchronies ¿ 0 and a regular logistic function for asynchronies ¡ 0, followed by applying a min-max normalization to the input asynchrony as illustrated in Equation 5 and Equation 6:

$$async_s = \begin{cases} t_{max} & abs(async) \geq t_{max} \\ \dfrac{2(abs(async) - t_{min})}{t_{max} - t_{min}} - 1 & t_{min} < abs(async) < t_{max} \\ t_{min} & abs(async) \leq t_{min} \end{cases} \tag{5}$$

$$comp_{time} = \begin{cases} \dfrac{2}{1 + e^{k_1(async_s)}} - 1 & async > 0 \\ 0 & async = 0 \\ \dfrac{2}{1 + e^{-k_2(async_s)}} - 1 & async < 0 \end{cases} \tag{6}$$

Here, $async$ is the time between events from the point of view of the finger tap, which is scaled to a value between -1 and +1, where t_{max} indicates a cut-off point in milliseconds, and t_{min} is equal to 0. Dependent on the sign of the $async$ this goes through a logistic function, where k_1 is the steepness of the curve for positive asynchronies, and k_2 for negative asynchronies. Thus,

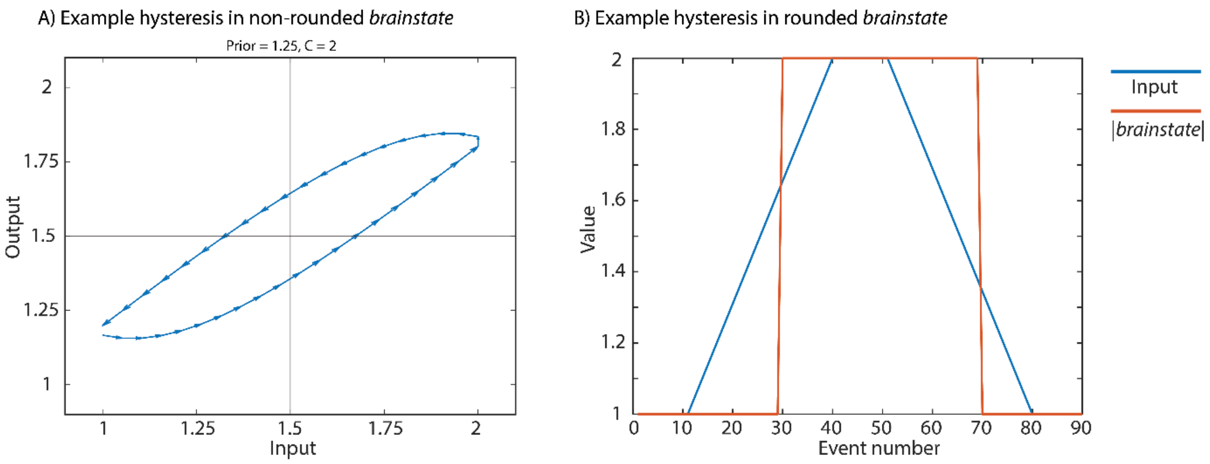
$comp_{time}$ approaches 1 when the finger tap and auditory feedback is considered linked, and approaches -1 when they are considered not linked.

$comp_{aud}$ is a comparison between auditory features, which calculates similarity between the expected sound caused by a finger tap, and what is heard. This is inherently a complex comparison, relying on multiple acoustic features, which we have left for further iterations of the MEAMSO. Hence, we here set $comp_{aud} = 0$ so it does not influence the rest of the calculations. As discussed in the main paper there are multiple perceptual models that could serve as a drop-in function in future expansions of the MEAMSO.

$comp_{nback}$ is an n-back instantaneous comparison, here a Bayesian average of the N previous $comp_{time}$. Its value is calculated as $comp_{nback} = \frac{Cm+\sum_{i=1}^{N} comp_{time_i}}{C+N}$, with a constant C and a prior m. This prior is updated for every calculation to be equal the previous value of $comp_{nback}$.

$comp_{corr}$ is a measure of the correlation between the time-series of finger taps and auditory events. We here calculate the cross-correlation function at lags 0 and -1, and select the maximum value so that $comp_{corr} = \max(lag_{-1}, lag_0)$.

Together this creates a model that, dependent on the values chosen for the free variables, exhibits hysteresis as illustrated in Supplementary Figure 1. A minimal implementation of the MEAMSO in MATLAB can be found at https://github.com/OleAd/MEAMSO. The demonstration script was written in MATLAB R2018b and uses the crosscorr.m function from the Econometrics Toolbox. However, we have included a fallback calculation if this function is not present.



Supplementary Figure 1 - In A we show an example of hysteresis in the non-rounded brainstate *shown in Equation 1. The input here is a vector that linearly goes from 1 to 2 over 20 steps, and then from 2 to 1 over 20 steps. The blue lines show the output at the given input with the blue arrows indicating the direction of the input vector. In B we show how the rounded* brainstate *value (in orange) looks for an example input, here shown in blue. This input repeats the value 1 for 10 steps, linearly increases to 2 over 30 steps, remains at 2 for 10 steps, before linearly decreasing to 1 over 30 steps, and finally remains at 1 for 10 steps. As can be seen, the* brainstate *remains in state 1 even though the input has passed the midpoint at 1.5, and do not switch to state 2 before the input is well above midpoint. A similar behaviour is shown for the opposite direction. These values crucially rely on the chosen prior* p *and constant* c *as shown in Equation 1, here* $\mathrm{p} = 1.25$ *and* $\mathrm{c} = 2$.