Supplemental Data

Dynamic associations in the cerebellar-motoneuron network during motor learning
Raudel Sánchez-Campusano, Agnès Gruart, and José M. Delgado-García

Supplementary Appendix S1
To quantify the dependence of a signal $Y_{EMG}(t)$ (for OO EMG record) on a signal $X_{NR}(t)$ (for IP or MN neuronal records), an estimator of the nonlinear correlation index ($\eta^2_{EMG|NR}$) was computed as follows (Meeren et al., 2002; Pereda et al., 2005; Ansari-Asl, et al., 2006; Kalitzin et al., 2007).

$$\eta^2_{EMG|NR} = 1 - \frac{\sum_{j=1}^{Nb} \sum_{k \in B_j} (Y_{EMG_k} - \bar{X}(X_{NR_j}))^2}{\sum_{k=1}^{Ns} (Y_{EMG_k} - \bar{Y}_{EMG})^2},$$

with $Ns$ being the number of samples of the electrophysiological signals and $\bar{Y}_{EMG}$ being the average of all $Y_{EMG_k}$. The measure of association in the opposite direction $\eta^2_{NR|EMG}$ can be calculated analogously. The equation for the linear piecewise approximation (LPA) of the nonlinear regression curve $\bar{X}(X_{NR})$, according to the $s_j$ binary sequence, is

$$\bar{X}(X_{NR}) = \sum_{j=1}^{Nb-1} s_j(X_{NR}) g_j(X_{NR}),$$

where

$$s_j = \begin{cases} 1 & \text{if } X_{NR} \leq p_2 \text{ else } 0, \text{ for } j = 1, \\ 1 & \text{if } p_j < X_{NR} \leq p_{j+1} \text{ else } 0, \text{ for } j = 2, ..., Nb - 2, \\ 1 & \text{if } X_{NR} > p_{Nb-1} \text{ else } 0, \text{ for } j = Nb - 1. \end{cases}$$

The equation for the successive segments of lines is

$$g_j(X_{NR}) = \frac{q_{j+1} - q_j}{p_{j+1} - p_j} (X_{NR} - p_j) + q_j,$$

In practice, a scatter plot of $Y_{EMG}(t)$ versus $X_{NR}(t)$ is studied (Figs. 2 and 3A, in the main text). The values of $X_{NR}(t)$ are subdivided into $Nb$ bins $B_j$, with $j = 1, ..., Nb$. For each bin, the $X_{NR}(t)$ value of the midpoint ($p_j$) and the
average value \( (q_j) \) of \( Y_{EMG}(t) \) are calculated. The curve of regression is approximated by connecting the resulting points \( (p_j, q_j) \) by segments of straight lines (according to \( g_j(X_{NR}) \), with \( j = 1, \ldots, Nb - 1 \)).

The nonlinear index \( \eta^2_{\{EMG|NR\}} \) ranges between 0, when both signals are independent \([Y_{EMG}(t) \text{ is completely independent of } X_{NR}(t)]\), and 1, for a perfect dependence \([Y_{EMG}(t) \text{ is fully determined by } X_{NR}(t)]\). In the case of a linear relationship between \( Y_{EMG}(t) \) and \( X_{NR}(t) \), the nonlinear index \( \eta^2 \) (with \( \eta^2_{\{EMG|NR\}} = \eta^2_{\{NR|EMG\}} \)) reduces to the common squared linear regression coefficient \( r^2 \). For a nonlinear relationship, \( \eta^2_{\{EMG|NR\}} \neq \eta^2_{\{NR|EMG\}} \), and the difference \( \Delta\eta^2 = \eta^2_{\{EMG|NR\}} - \eta^2_{\{NR|EMG\}} \) indicates the degree of asymmetry of the nonlinear coupling between the electrophysiological records \( X_{NR}(t) \) and \( Y_{EMG}(t) \) (Pereda et al., 2005).

By calculating the index \( \eta^2 \), it is also possible to estimate the delay in the coupling between the signals. Similarly, as in the case of the cross-correlation function, \( \eta^2 \) can be estimated as a function of time shift \( (\tau) \) between signal \( X_{NR}(t) \) and \( Y_{EMG}(t) \) or vice versa. That shift for which the maximum value for \( \eta^2(\tau) \) is reached is used as an estimate of the time lag (or time delay) between the two signals. Indeed, if \( X_{NR}(t) \) causes \( Y_{EMG}(t) \), \( \tau_{\{EMG|NR\}} \) (corresponding to \( \eta^2_{\{EMG|NR\}} \)) will be positive and \( \tau_{\{NR|EMG\}} \) (corresponding to \( \eta^2_{\{NR|EMG\}} \)) will be negative, so that the difference \( \Delta\tau = \tau_{\{EMG|NR\}} - \tau_{\{NR|EMG\}} \) will also be positive.

By combining the information of asymmetry and of time of delay in coupling, the following direction index has been proposed (Wendling et al., 2001) as providing a robust measure of the direction of coupling:

\[
D_\pm = \frac{1}{2} [\text{sgn}(\Delta\eta^2) + \text{sgn}(\Delta\tau)],
\]

\( D_\pm = +1 \), indicates a unidirectional coupling \([X_{NR}(t) \overset{YES}{\longrightarrow} Y_{EMG}(t)]\) from signal \( X_{NR}(t) \) (for VII n and IP n nuclei) to signal \( Y_{EMG}(t) \) (for the motor units).
$D_{\pm} = -1$, indicates a unidirectional \( Y_{EMG}(t) \rightarrow X_{NR}(t) \) coupling from signal \( Y_{EMG}(t) \) (for the motor units) to signal \( X_{NR}(t) \) (for VII n and IP n nuclei).

$D_{\pm} = 0$, indicates a bidirectional \( Y_{EMG}(t) \leftrightarrow X_{NR}(t) \) coupling between the two signals (i.e., a “trivial” feedback relationship in the network association).

On the other hand, we can quantify whether a change in some control variable significantly alters an existing nonlinear correlation between two other variables. In other words, we can decide if the values of the nonlinear association indices \( \eta \), see Fig. 3B in the main text) are significantly high or low. This is done by using modified Fisher’s z-transformation \( w \) to associate each measured nonlinear association index \( \eta \) with a corresponding w-transformation. Thus, \( \eta \) can be normalized according to the following equation:

\[
\eta = \frac{1}{2} \ln \left( \frac{1 + \eta}{1 - \eta} \right),
\]

a transformation that modified the conventional z-transformation

\[
z = \frac{1}{2} \ln \left( \frac{1 + \eta}{1 - \eta} \right)
\]

with values in \([0, +\infty)\). Then, each \( w \)-transformation is approximately normally distributed with values in \([-\infty, +\infty]\). If the probability \( p \) is less than the predetermined significance level (0.05), then reject the null hypothesis \( H_0 \), no correlation). A small value of \( p \) indicates that the two distributions \( w \)-transformation and Gaussian) are significantly correlated. The confidence bounds are based on an asymptotic normal distribution of \( w \)-transformation, with a mean \( \bar{w} \) and a standard deviation \( \text{std}(w_{NR|EMG}) \). These bounds are accurate for large samples when the data has a multivariate normal distribution. The \( p \) values were calculated using the complementary error functions (erfc) for modified Fisher’s \( w \)-transformations of the nonlinear association indices.

\[
p[w_{NR|EMG}] = k \cdot \text{erfc} \left( \frac{|w_{NR|EMG} - \bar{w}|}{\sqrt{2} \cdot \text{std}(w_{NR|EMG})} \right)
\]

For the nonlinear association curves in Figure 3B, the null hypothesis \( (H_0, \text{ no correlation}) \) were rejected for all the \( w \)-transformations (modified Fisher’s z-tests, \( H_0 = 1 \); \( p[w_{O_E|EMG} | MNR] < 0.01 \), \( p[w_{MNR|O_E|EMG}] < 0.01 \),
\[ p[w_{\text{MN NR}\mid\text{EMG}}] < 0.01 \quad \text{and} \quad p[w_{\text{IP NR}\mid\text{EMG}}] < 0.05 \] of the nonlinear association indices (\( \eta \)).

Similarly, the significance of a difference (\( H_0 \), no difference) between two measured nonlinear correlation coefficient (e.g., \( \eta_{\text{MN NR}\mid\text{EMG}} \) and \( \eta_{\text{IP NR}\mid\text{EMG}} \); \( \eta_{\text{MN NR}\mid\text{EMG}} \) and \( \eta_{\text{IP NR}\mid\text{EMG}} \)) was calculated using their corresponding w-transformations and complementary error functions for pairs of means samples, with variances \( \text{var}(w_{\text{MN NR}\mid\text{EMG}}) \) and \( \text{var}(w_{\text{IP NR}\mid\text{EMG}}) \), respectively (modified Fisher’s z-tests, \( H_0 = 1 \),
\[ p[w_{\text{MN NR}\mid\text{EMG}}] < 0.05 \], for \( w_{\text{MN NR}\mid\text{EMG}} \) and \( w_{\text{IP NR}\mid\text{EMG}} \);
\( H_0 = 1 \), \( p[w_{\text{MN NR}\mid\text{EMG}}] < 0.05 \), for \( w_{\text{MN NR}\mid\text{EMG}} \) and \( w_{\text{IP NR}\mid\text{EMG}} \).

As a consequence the ANOVA test revealed the results illustrated in the Figure S1 for the nonlinear association curves in the Figure 3B of the main text. The \( p \)-value indicates that the four w-transformations of the nonlinear association indices were significantly different and \( H_0 \) (no difference between population means) was rejected, then the alternative working hypothesis must be true (ANOVA F-test, \( H_0 = 1 \); \( F_{(3,9,2000)} = 1844.35 \), \( p < 0.01 \), see Table S1).

Sometimes it is preferable to perform a test to determine which pairs of means are significantly different, and which are not. In a one-way ANOVA, we compare the means of several groups (i.e., the w-transformations of the nonlinear association indices) to test the hypothesis that they are all the same, against the general alternative that they are not all the same. Sometimes this alternative may be too general. We may need information about which pairs of means are significantly different, and which are not (according to nonlinear association indices and their w-transformations, Fig. 3B in the main text). A test that can provide such information is called a "multiple comparison". Multiple comparison procedures were designed to provide an upper bound on the probability that any comparison will be incorrectly found significant (see, Fig. S2 and Table S2).
Supplementary Appendix S2

In practice, physiological time series, such as $X_{\text{NR} MN}(t)$, $X_{\text{NR} IP}(t)$, $Y_{\text{EMG} \text{OFF}}(t)$, $f_{\text{MN}}(t)$, $f_{\text{IP}}(t)$ and $\theta(t)$ exhibit non-stationary behavior. To reach stationarity in the models, the cumulative areas underneath the instantaneous frequency functions and the learned motor responses (or integrated neuronal and eyelid activities) were calculated, and High-Pass Filters (HPF) were designed to estimate the relative variation function during the learning process (e.g., Fig. 6, in the main text). Additionally, simultaneous differencing of all resultant relative variation functions avoided unnecessary complications in models fitting, and — as a consequence — the stationary time series for the transfer function models were obtained: $S0_t(\theta)$ for $\theta(t)$, $S1_t(MN)$ for $f_{\text{MN}}(t)$, and $S2_t(IP)$ for $f_{\text{IP}}(t)$.

Thus, the linear combinations of the elements of these physiological time series may often be stationary. When several such series are considered jointly, non-stationarity may be modeled by allowing the zeros of the determinantal polynomial corresponding to the autoregressive operator to lie on the unit circle. The time series will be stationary and invertible simultaneously when the zeros of the determinantal polynomials, corresponding to the autoregressive

$$\delta_a(B) = 1 - \sum_{k_a = 1}^{a} \delta_{k_a} B^{k_a}$$

and moving average $\omega_m(B) = \sum_{k_m = 0}^{m} \omega_{k_m} B^{k_m}$ operators, are all outside the unit circle (Tiao and Box, 1981). The sample impulsive response

$$\nu(B) = \frac{\omega_m(B)}{\delta_a(B)} B^{b} = \sum_{k = 0}^{\nu_k} B^{k}$$

is also a function of $k$ (scalar integer indicating the number of lags of the impulsive response to compute), where the parameters $\{\nu_k\}$ are the transfer function model coefficients.

Two of the reasons for analyzing and modeling such physiological time series jointly are:

- a) To understand the dynamic relationships between them. They may be contemporaneously related, one series may lead the others, or there may be feedback relationships.
- b) To improve the accuracy of forecasts. When there is information on one series contained in the historical data of another, better forecasts can result when the series are modeled jointly.
At the same time, three basic principles for the construction of a transfer function model are:

a) The relationship between the input time series $S_I_t$ and the output dependent time series $S_O_t$ is constant in the sample period of analysis. The response $S_O^*_{t}$ to variations of $S_I_t$ can be approximated according to the linear (or locally linear) formulation of the form 

$$S_O^*_{t} = \sum_{k=0}^{\infty} v_k S_I_{t-k} = v(B) S_I_t.$$ 

b) The input time series $S_I_{t-k}$ (or $S_I_t$) affects the output dependent time series $S_O_t$ (or $S_O_{t+k}$), for $k \geq 0$. In contrast, there is no relationship between $S_O_{t-k}$ (or $S_O_t$) and $S_I_t$ (or $S_I_{t+k}$) for $k > 0$. This condition demands a unidirectional coupling (without feedback) between the time series.

c) The input time series $S_I_t$ affects the output dependent time series $S_O_t$ and vice versa, for both $k > 0$ and $k < 0$: there is a relationship between $S_O_{t-k}$ (or $S_O_t$) and $S_I_t$ (or $S_I_{t+k}$) for $k > 0$; and between $S_I_{t+k}$ (or $S_I_t$) and $S_O_t$ (or $S_O_{t-k}$) for $k < 0$). In this condition, both the input time series $S_I_t$ and the output time series $S_O_t$ can have feedback relationships (this indicates a bidirectional coupling between the time series).

In the particular case when the period $b = 0$, then the transmission of the effects between the time series is instantaneous, and the transfer function model is simplified to an input-output system, where the current output $S_O^*_{t}$ is the linear combination of previous $a$-outputs, the current input $S_I_t$, and the previous $m$-inputs.

$$S_O^*_{t} = \sum_{k_a=1}^{a} \delta_{k_a} B^{k_a} S_O^*_{t} + \sum_{k_m=0}^{m} \omega_{k_m} B^{k_m} S_I_t = \sum_{k_a=1}^{a} \delta_{k_a} S_O^*_{t-k_a} + \sum_{k_m=0}^{m} \omega_{k_m} S_I_{t-k_m}$$

Considering the process of inertia (or uncertainties, $U_t$) of the transfer function model $S_O_t = S_O^*_{t} + U_t$, equaling the orders of the autoregressive and
moving average operators \((s = a = m)\), and leaving the dependence between
the current output \(S0^*_t\) and the current input \(SI_t\) without effect, the following
bivariate autoregressive model is obtained:

\[
S0_t = \sum_{k=1}^{s} \delta^0_k S0_{t-k} + \sum_{k=1}^{s} \omega^0_k SI_{t-k} + u^0_t, \\
SI_t = \sum_{k=1}^{s} \delta^l_k SI_{t-k} + \sum_{k=1}^{s} \omega^l_k S0_{t-k} + u^l_t
\]

where \(\delta^0_k\), \(\omega^0_k\), \(\delta^l_k\) and \(\omega^l_k\) are the parameters of the model, \(s\) is the model
order, and \(u^0_t\) and \(u^l_t\) are the uncertainties or residual noises associated with
the model \((u^0_t\) is the prediction error when \(S0_t\) is predicted from its own past
and the past of \(SI_t\), and similarly for \(u^l_t\); i.e., the prediction error for each
individual time series depends on the past values of the two time series).

On the other hand, for the univariate model

\[
S0_t = \sum_{k=1}^{s} \delta^0_k S0_{t-k} + u^0_t, \text{ univariate autoregressive equation for } S0_t \\
SI_t = \sum_{k=1}^{s} \delta^l_k SI_{t-k} + u^l_t, \text{ univariate autoregressive equation for } SI_t
\]

where the parameters \(\delta^0_k\), and \(\delta^l_k\) are the model coefficients, and \(u^0_t\) and \(u^l_t\)
are the respective prediction errors when \(S0_t\) and \(SI_t\) are predicted from their
own past.

Coefficients \(\delta^0_k\), \(\delta^l_k\), \(\delta^0_{10}\), \(\omega^0_k\), \(\delta^l_{10}\) and \(\omega^l_k\) are selected in order to
minimize the mean squared error for both univariate \((\varepsilon^2_0\) and \(\varepsilon^2_l\)\) and bivariate
\((\varepsilon^2_{10}\) and \(\varepsilon^2_{0l}\)\) models:

\[
\varepsilon^2_0 = \frac{1}{N} \sum \left[ S0_t - \sum_{k=1}^{s} \delta^0_k S0_{t-k} \right], \text{ and } \varepsilon^2_l = \frac{1}{N} \sum \left[ SI_t - \sum_{k=1}^{s} \delta^l_k SI_{t-k} \right],
\]
\[ \varepsilon_{I0}^2 = \frac{1}{N} \sum \left[ S0_t - \left( \sum_{k=1}^{s} \delta_{k} S0_{t-k} + \sum_{k=1}^{s} \omega_{k} S1_{t-k} \right) \right], \]

where \( N \) is the number of predicted samples in a time series. If the number of model parameters is much smaller than the number of data points used for the estimation, minimal values of these errors (\( \varepsilon_{0}^2, \varepsilon_{I}^2, \varepsilon_{I0}^2 \) and for the opposite direction \( \varepsilon_{0I}^2 \)) by an equation similar to \( \varepsilon_{I0}^2 \) in which \( I \) and \( 0 \) should be interchanged) can characterize the accuracy of the models: the smaller the error, the better the model (self-predictability of the model).

The optimal model order (\( s \)) is determined according to the criteria usually based on the statistics constructed from prediction errors. Akaike’s Information Criterion (AIC) is the most commonly used one:

\[ AIC(j) = N \log(\text{det}(CM_j)) + 2L^2j, \]

where \( CM_j \) is the estimate of the prediction error covariance matrix assuming a \( j \)th order model, \( N \) is the number of data points and \( L \) is the number of variables. For reliable results, \( AIC(j) \) should be computed only for a maximum value of \( j \) of \( 3\sqrt{N}/L \). The optimal model order (\( s \)) corresponds to the minimum of \( AIC(j) \).

The prediction performance for the two models (bivariate and univariate) can be assessed by the variances of the prediction errors: \( V_{I0} = \text{var}(u_{I0}^t) \), \( V_{0I} = \text{var}(u_{0I}^t) \), \( V_0 = \text{var}(u_0^1) \) and \( V_I = \text{var}(u_I^1) \), where \( \text{var}(\cdot) \) indicates variance operator. The covariance matrix of \( u_{I0}^t \) and \( u_{0I}^t \) is of the form

\[
CM = \begin{bmatrix}
V_{I0} & V_{I0|0I} \\
V_{0I|I0} & V_{0I}
\end{bmatrix},
\]

where \( V_{I0|0I} \) and \( V_{0I|I0} \) are the off-diagonal elements, and represent the mixed covariances of the prediction error vector.

The measurement of time-dependent Granger causality from \( S1_t \) and \( S0_t \) stationary time series is defined as \( G_{I \rightarrow 0} = \ln(V_0/V_{I0}) \). If the past of \( S1_t \) does not improve the prediction of \( S0_t \), then \( V_{I0} \approx V_0 \) and causality measurement will be close to zero \( G_{I \rightarrow 0} = 0 \), which indicates \( S1_t \) does not
cause $S_0_t$. If $V_{I0} < V_0$, then $SI_t$ causes $S0_t$ in the sense of Granger causality (Geweke, 1982; Nolte et al., 2008). Any improvement in the prediction of $S0_t$ by the inclusion of $SI_t$ leads to a decrease in $V_{I0}$, thereby increasing the causality. The measurement of $G_{I \rightarrow 0}$ is invariant when $SI_t$ and $S0_t$ are rescaled or pre-multiplied by different invertible lag operators. Symmetrically, the measurement of Granger causality for the opposite direction, from $S0_t$ to $SI_t$, is defined as $G_{0 \rightarrow I} = \ln(V_I / V_{0I})$. If both $G_{I \rightarrow 0}$ and $G_{0 \rightarrow I}$ are high, this indicates a bidirectional coupling or feedback relationship between the time series (see Fig. 5, in the main text).

The causality index $G_{I \rightarrow 0}$ is nonnegative and measures the reduction in the total variance of predictive errors of $S0_t$ when past $SI_t$ is added for prediction. The percentage reduction of the variance suggests the degree of $S0_t$ relating to the history of $SI_t$, and vice versa. Finally, the normalized causality indexes are given as $R_{I \rightarrow 0}^2 = 1 - e^{-G_{I \rightarrow 0}}$ and $R_{0 \rightarrow I}^2 = 1 - e^{-G_{0 \rightarrow I}}$.

In order to measure coupling between the input $SI_t$ and output $S0_t$ time series, we may also use the normalized prediction improvement, the so-called Granger-Sargent statistic:

$$GS_{I \rightarrow 0}^2 = \frac{\varepsilon_{0I}^2 - \varepsilon_{I0}^2}{\varepsilon_{I0}^2}, \text{ and for the opposite direction: } GS_{0 \rightarrow I}^2 = \frac{\varepsilon_{I0}^2 - \varepsilon_{0I}^2}{\varepsilon_{0I}^2},$$

This method examines mean squared prediction errors, $\varepsilon_{0I}^2$ or $\varepsilon_{I0}^2$, and when these errors are smaller than $\varepsilon_{0I}^2$ or $\varepsilon_{I0}^2$, respectively, then it is assumed that process $SI_t$ affects process $S0_t$ (according to the positive value of $GS_{I \rightarrow 0}^2$) or vice versa (according to the positive value of $GS_{0 \rightarrow I}^2$).
Supplementary References


Supplementary Figure Legends

Figure S1. The ANOVA test and the box plot for the nonlinear association curves of the Figure 3B in the main text and their w-transformations. A–B, The very small p-values of 3.1x10^{-15} (case 1, A) and 2.7x10^{-15} (case 2, B) indicate that differences between the w-group means are highly significant. The probability of this outcome under the null hypothesis (i.e., the probability that samples actually drawn from the same population would have means differing by the amounts seen in data) is less than 31 in 10^{14} (case 1, A) and 27 in 10^{14} (case 2, B). The tests therefore strongly support the alternate hypothesis, that one or more of the samples are drawn from populations with different means (one-way ANOVA F-tests, H0=1; F_{(3,9,2000)}=1844.35, p < 0.01; and H0=1, F_{(3,9,2000)}=1845.35, p < 0.01) and the box plots confirm this graphically. The box has lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the box to show the extent of the rest of the data. A, In the case 1, the estimated values of the medians were: 0.7783 for w_{(OO EMG|MN NR)}, 0.7181 for w_{(MN NR|OO EMG)}, -0.2320 for w_{(OO EMG|IP NR)} and -0.2370 for w_{(IP NR|OO EMG)}. B, In the case 2, the estimated values of the medians were: 0.3816 for w_{(OO EMG|MN NR)}, 0.3153 for w_{(MN NR|OO EMG)}, -0.8711 for w_{(OO EMG|IP NR)} and -0.8795 for w_{(IP NR|OO EMG)} (see, the ANOVA Table S1).

Figure S2. The multiple comparison procedure for the nonlinear association curves of the Figure 3B in the main text and their w-transformations. The multiple comparison graph and the quantitative analysis (rows of the output matrix in Table S2) shows that all of the comparisons involving the four groups (w_{(OO EMG|MN NR)}, w_{(MN NR|OO EMG)}, w_{(OO EMG|IP NR)} and w_{(IP NR|OO EMG)}) have confidence intervals that do not include zero, so the difference is significant at the 0.05 level (a 95% confidence interval for the true estimated difference in means). In other words, we can reject the null hypothesis that the true difference is zero and therefore, those differences are significant. A, In the case 1, the estimated values of the means were: 0.8114 for w_{(OO EMG|MN NR)}, 0.7273
for \( w_{\langle \text{MN NR} | \text{OO EMG} \rangle} \), -0.2195 for \( w_{\langle \text{OO EMG} | \text{IP NR} \rangle} \) and -0.2982 for \( w_{\langle \text{IP NR} | \text{OO EMG} \rangle} \) (see the confidence intervals for the difference in means in Table S2). B, In the case 2, the estimated values of the means were: 0.4076 for \( w_{\langle \text{OO EMG} | \text{MN NR} \rangle} \), 0.3165 for \( w_{\langle \text{MN NR} | \text{OO EMG} \rangle} \), -0.8828 for \( w_{\langle \text{OO EMG} | \text{IP NR} \rangle} \) and -0.9775 for \( w_{\langle \text{IP NR} | \text{OO EMG} \rangle} \). Two means are significantly different if their intervals are disjoint, and are not significantly different if their intervals overlap. The standard errors of the means (SEM) were 0.0139 (case 1, A) and 0.0174 (case 2, B).

**Figure S3.** Waveform analysis of conditioned eyelid responses (CR). A, Illustration of an eyelid CR (position, in degrees) collected from a single trial during the 10th conditioning session. The onset of the CR (red arrow, according to the preceding eyelid position: see Fig. 1K in the main text) was determined as the latency corresponding to the interception (the red circle) of the regression function (see, e1, regression line) with the maximum amplitude level (red dashed line). Also, the action potentials (IP spikes), marked with blue plus signs, correspond to the direct representation of the firing activity in the IP nucleus neuron collected during the same trial. B, High-Pass Filtering (HPF) of the integrated eyelid activity (see Fig. 6B in the main text). C, Fast Fourier Transforms (FFT). Power spectra for the oscillating curves shown in B. Note that the two illustrated spectra (black line and magenta circles), during the 10th conditioning session, presented a significant predominance of spectral components at \( \approx 20 \) Hz.
Supplementary Tables

Table S1. ANOVA parameters as a result of the statistic analysis for the \( w \)-transformations of the nonlinear association indices (\( \eta \), see Fig. 3B in the main text and Fig. S1 in this supplementary material) during the 10th conditioning session.

<table>
<thead>
<tr>
<th>ANOVA Test</th>
<th>CASE 1 (Fig. S1A)</th>
<th>CASE 2 (Fig. S1B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>df</td>
<td>SS</td>
</tr>
<tr>
<td>w-groups</td>
<td>3</td>
<td>532.96</td>
</tr>
<tr>
<td>Error</td>
<td>2000</td>
<td>192.65</td>
</tr>
<tr>
<td>Total</td>
<td>2003</td>
<td>725.61</td>
</tr>
</tbody>
</table>

Table S2. The characteristic values of the multiple comparisons algorithm for the \( w \)-transformations (groups) of the nonlinear association indices (\( \eta \), see Fig. 3B in the main text and Fig. S2 in this supplementary material) during the 10th conditioning session.

<table>
<thead>
<tr>
<th>CASE 1 (Fig. S2A)</th>
<th>A 99 % confidence interval for the difference in means</th>
<th>CASE 2 (Fig. S2B)</th>
<th>A 95 % confidence interval for the difference in means</th>
</tr>
</thead>
<tbody>
<tr>
<td>w-transformations</td>
<td>Smallest value</td>
<td>The estimated difference in means</td>
<td>Largest value</td>
</tr>
<tr>
<td>1 2</td>
<td>0.0231</td>
<td>0.0841</td>
<td>0.1452</td>
</tr>
<tr>
<td>1 3</td>
<td>0.9698</td>
<td>1.0309</td>
<td>1.0919</td>
</tr>
<tr>
<td>1 4</td>
<td>0.0486</td>
<td>1.1096</td>
<td>1.1707</td>
</tr>
<tr>
<td>2 3</td>
<td>0.8857</td>
<td>0.9467</td>
<td>1.0078</td>
</tr>
<tr>
<td>2 4</td>
<td>0.9645</td>
<td>1.0255</td>
<td>1.0866</td>
</tr>
<tr>
<td>3 4</td>
<td>0.0177</td>
<td>0.0788</td>
<td>0.1398</td>
</tr>
</tbody>
</table>