Supplementary Information

Accumulation of virtual tokens towards a jackpot reward enhances performance and value encoding in dorsal anterior cingulate cortex

Demetrio Ferro[1,2], Habiba Azab[3], Benjamin Hayden[3] and Rubén Moreno-Bote[1,2,4]

[1] Center for Brain and Cognition, Universitat Pompeu Fabra, 08002, Barcelona, Spain;

[2] Department of Information and Communications Technologies, Universitat Pompeu Fabra, 08002, Barcelona, Spain;

[4] Department of Neurosurgery, Baylor College of Medicine, Houston, TX, 77030;

[5] Serra Húnter Fellow Programme, Universitat Pompeu Fabra, Barcelona, 08002, Spain.

*Corresponding author: demetrio.ferro@upf.edu

	No Jackpot $JPT = 0$	Jackpot $JPT = 1$	Low $ATC = [0, 1]$	Medium $ATC = [2, 3]$	High $ATC = [4, 5]$
Subject 1	38606 (82.09%)	8424 (17.91%)	20674 (43.96%)	15222 (32.37%)	11134 (23.67%)
Subject 2	48910 (83.02%)	10001 (16.98%)	26889 (45.64%)	19057 (32.35%)	12965 (22.01%)

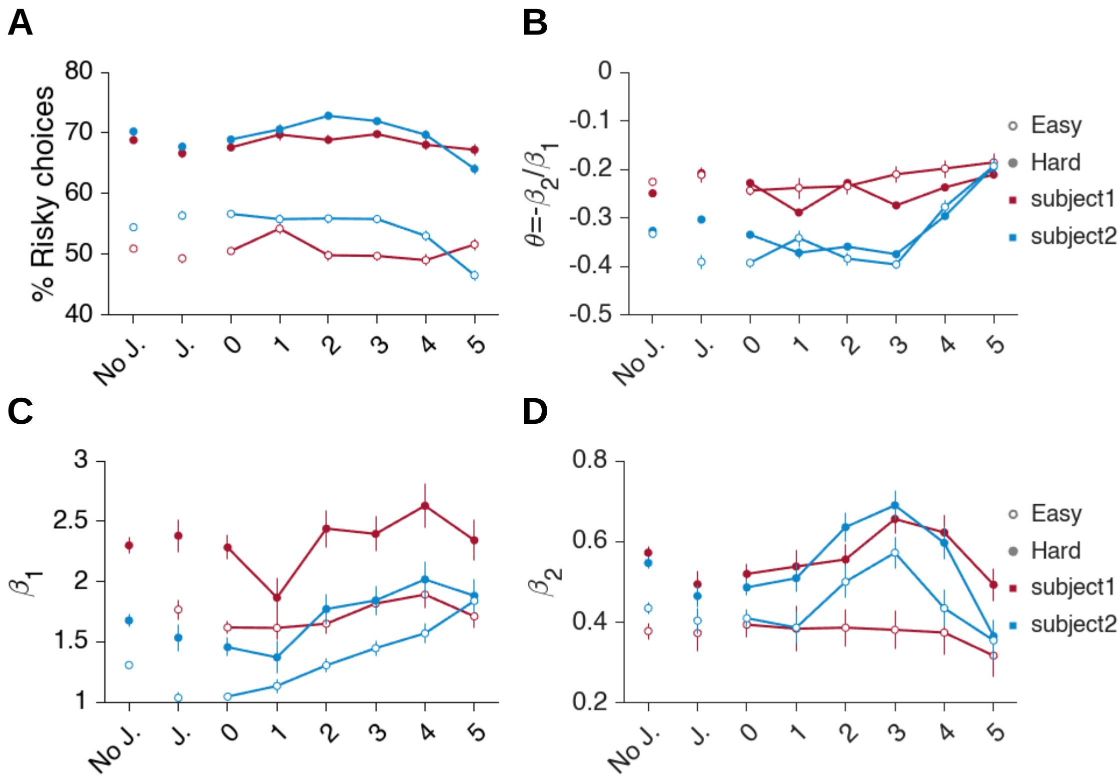
	No Jackpot $JPT = 0$	Jackpot $JPT = 1$	Low $ATC = [0, 1]$	Medium $ATC = [2, 3]$	High $ATC = [4, 5]$
Subject 1 Easy	19504 (41.57%)	4170 (9.05%)	10289 (22.08%)	7388 (16.66%)	5547 (11.88%)
Hard	19552 (40.51%)	4254 (8.87%)	10385 (21.88%)	7834 (15.71%)	5587 (11.79%)
Subject 2 Easy	24628 (41.22%)	4977 (8.53%)	13588 (22.58%)	9500 (16.22%)	6517 (10.95%)
Hard	24282 (41.81%)	5024 (8.45%)	13301 (23.07%)	9557 (16.13%)	6448 (11.06%)

Supplementary Table ST1: Trials available for jackpot and accumulated token counts ranges. Numbers of trials (fractions of the total) available for different token counts indexed by JPT and ATC ranges. The variable JPT indicates the presence ($JPT = 1$) or absence ($JPT = 0$) of jackpot reward on previous trial. The variable ATC is the accumulated tokens count as of the start of the current trial. The number of trials and fractions of the total are also shown for Easy vs Hard trials. The median value of Δ_{EV} is $median(\Delta_{EV}) = 1$ and is used to split data in 'Hard' and 'Easy' trials in behavioral analyses. These ATC settings are used in behavioral analyses. In neural analyses, we use 'Low' ($ATC = [0, 1]$) and 'High' ($ATC = [2, 5]$) as this allows us to improve the trial availability per condition.

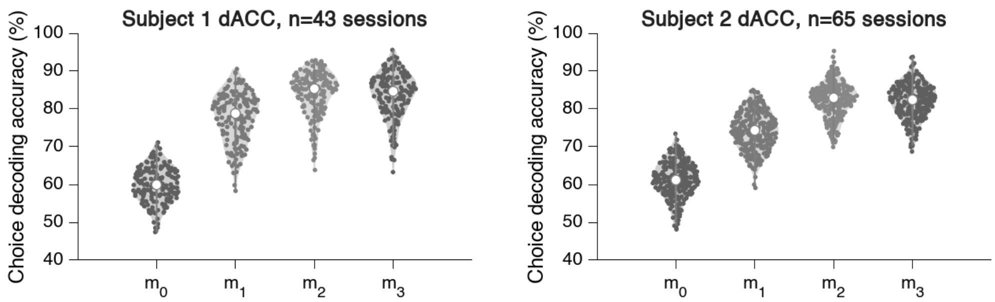
	$\Delta_{EV} \leq 0.5$	$0.5 < \Delta_{EV} \leq 1.5$	$1.5 < \Delta_{EV} \leq 2.5$	$2.5 < \Delta_{EV} \leq 3.5$	$\Delta_{EV} > 3.5$
Subject 1	13672 (29.07%)	19412 (41.28%)	10282 (21.86%)	3235 (6.88%)	429 (0.91%)
Subject 2	17152 (29.12%)	23945 (40.65%)	13298 (22.57%)	3996 (6.78%)	520 (0.88%)

	Hard: ($\Delta_{EV} < 1$)	Easy: ($\Delta_{EV} \geq 1$)
Subject 1	23806 (50.62%)	23224 (49.38%)
Subject 2	29306 (49.75%)	29605 (50.25%)

Supplementary Table ST2: Trials available for different difficulty levels.
Numbers of trials (fractions of the total) available for different difficulty levels indexed by Δ_{EV} ranges. Easier trials have larger Δ_{EV}. The median value of Δ_{EV} is $median(\Delta_{EV}) = 1$ and is used to split data in 'Hard' and 'Easy' trials in behavioral analyses. Note that in neural analyses we use Δ_{SV} based on SV.

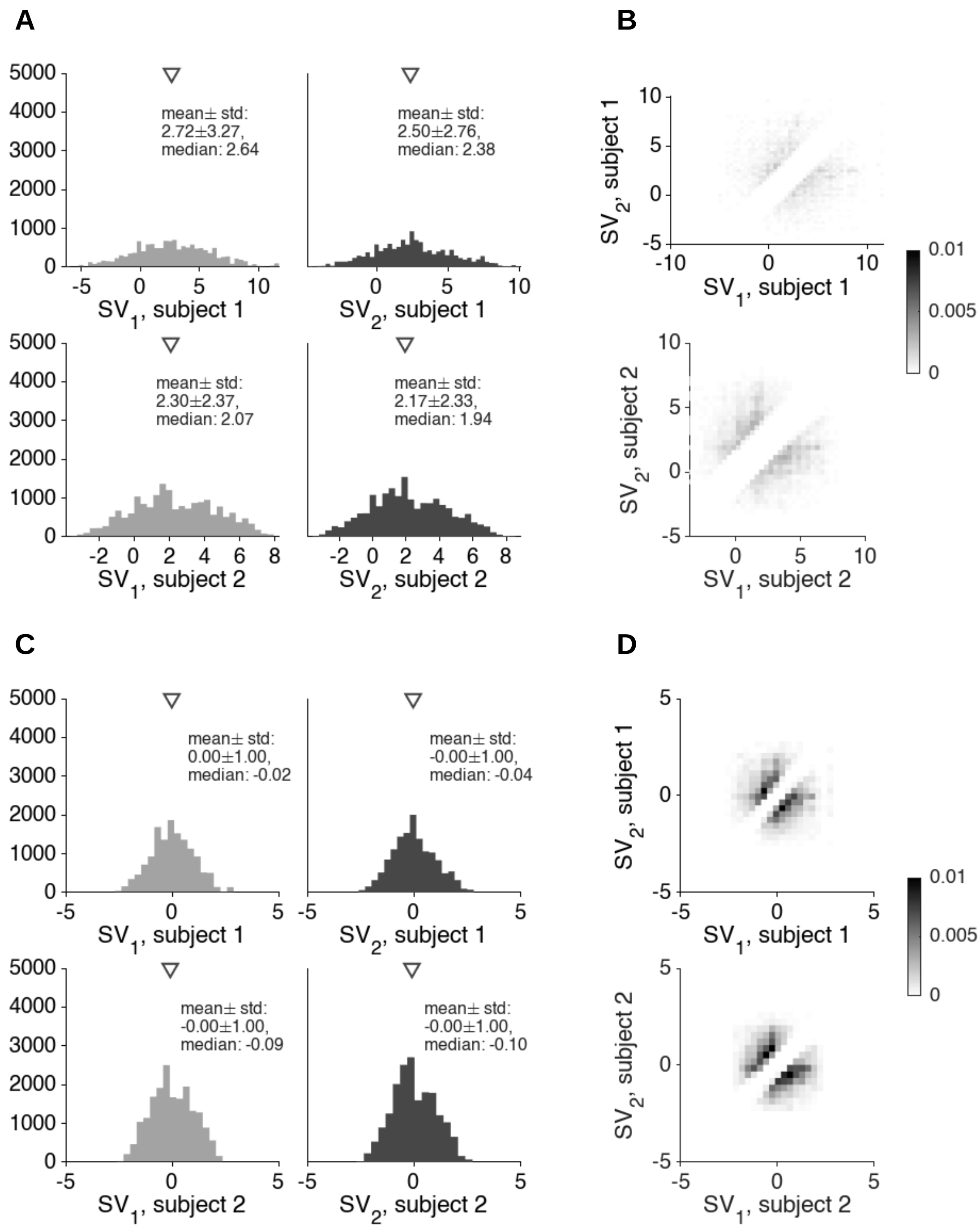


Supplementary Figure S1. Risk seeking attitude vs accumulated reward and difficulty. **A)** Fraction of trials (mean ± s.e.m.) with choice for the offer with higher risk. The data are split in 'No Jackpot' ($JPT = 0$), 'Jackpot' ($JPT = 1$), and for accumulated tokens count ($ATC = 0 - 5$). Data are split for subject 1 (red) and subject 2 (blue), and for Easy ($\Delta_{EV} \geq 1$, empty markers) and Hard ($\Delta_{EV} < 1$, filled markers). **B)** Markowitz risk return model for the offer utility based on the mean value (EV) and risk (R) of the offers. The model parameter (θ) describes risk attitude ($\theta < 0$ risk seeking, $\theta > 0$ risk avoiding) for jackpot cases and for values of accumulated token counts. Data is split in difficulty and across subjects as in A. **C)** Markowitz risk return model, parameter β_1 relative to EV difference weights for jackpot cases and for values of accumulated token counts. Data is split as in A. **D)** Markowitz risk return model, parameter β_2 relative to risk R difference weights for jackpot cases and for values of accumulated token counts. Data is split as in A.

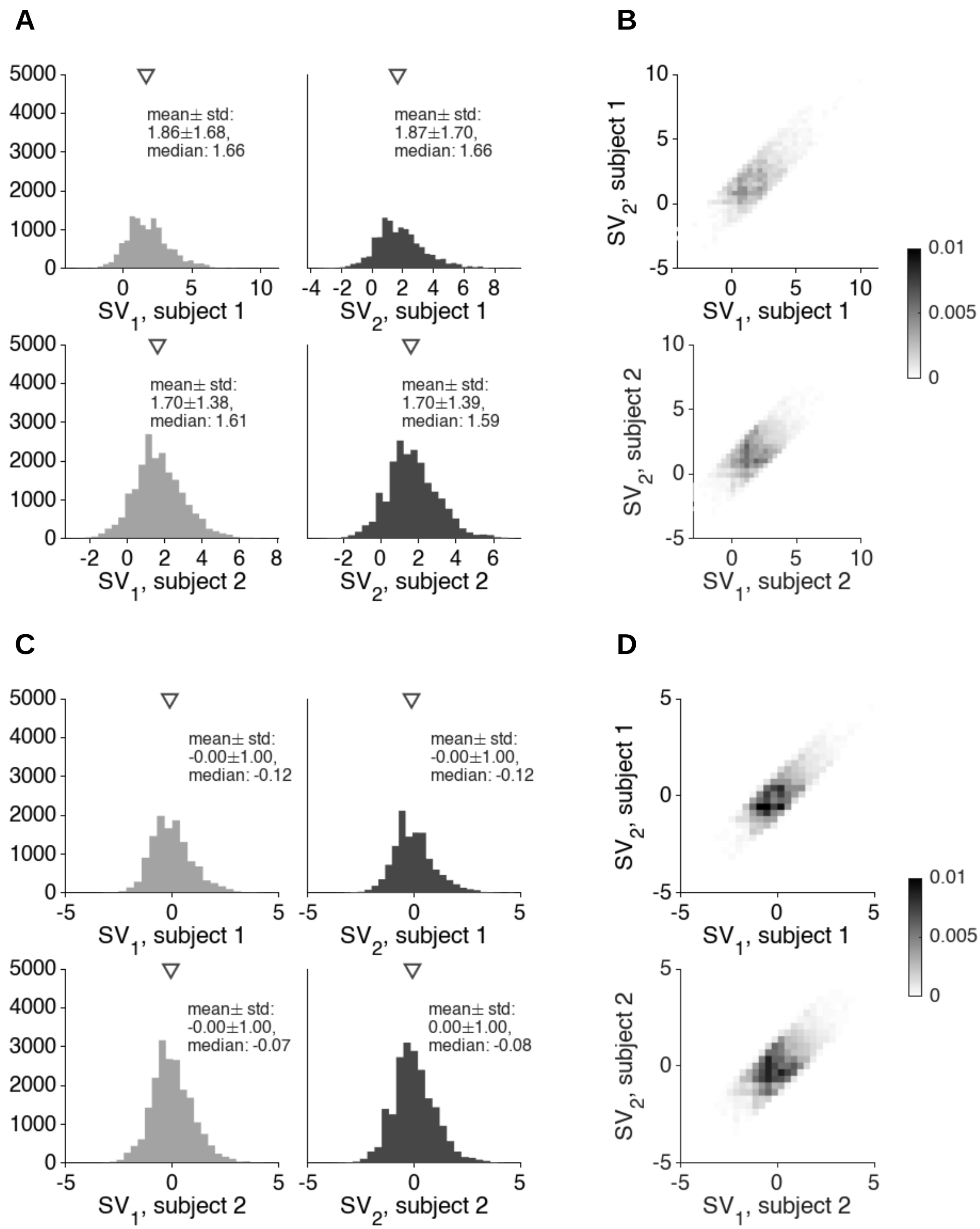


m_0: $\text{logit}(ch1) = \beta_0 + \beta_1 v_1^t + \beta_2 v_1^b - \beta_3 v_2^t - \beta_4 v_2^b$;
m_1: $\text{logit}(ch1) = \beta_0 + \beta_1 EV_1 - \beta_2 EV_2$;
m_2: $\text{logit}(ch1) = \beta_0 + \beta_1 EV_1 - \beta_2 EV_2 + \beta_3 R_1 - \beta_4 R_2$;
m_3: $\text{logit}(ch1) = \beta_0 + \beta_1 v_1^t + \beta_2 v_1^b - \beta_3 v_2^t - \beta_4 v_2^b + \beta_5 EV_1 - \beta_6 EV_2 + \beta_7 R_1 - \beta_8 R_2$.

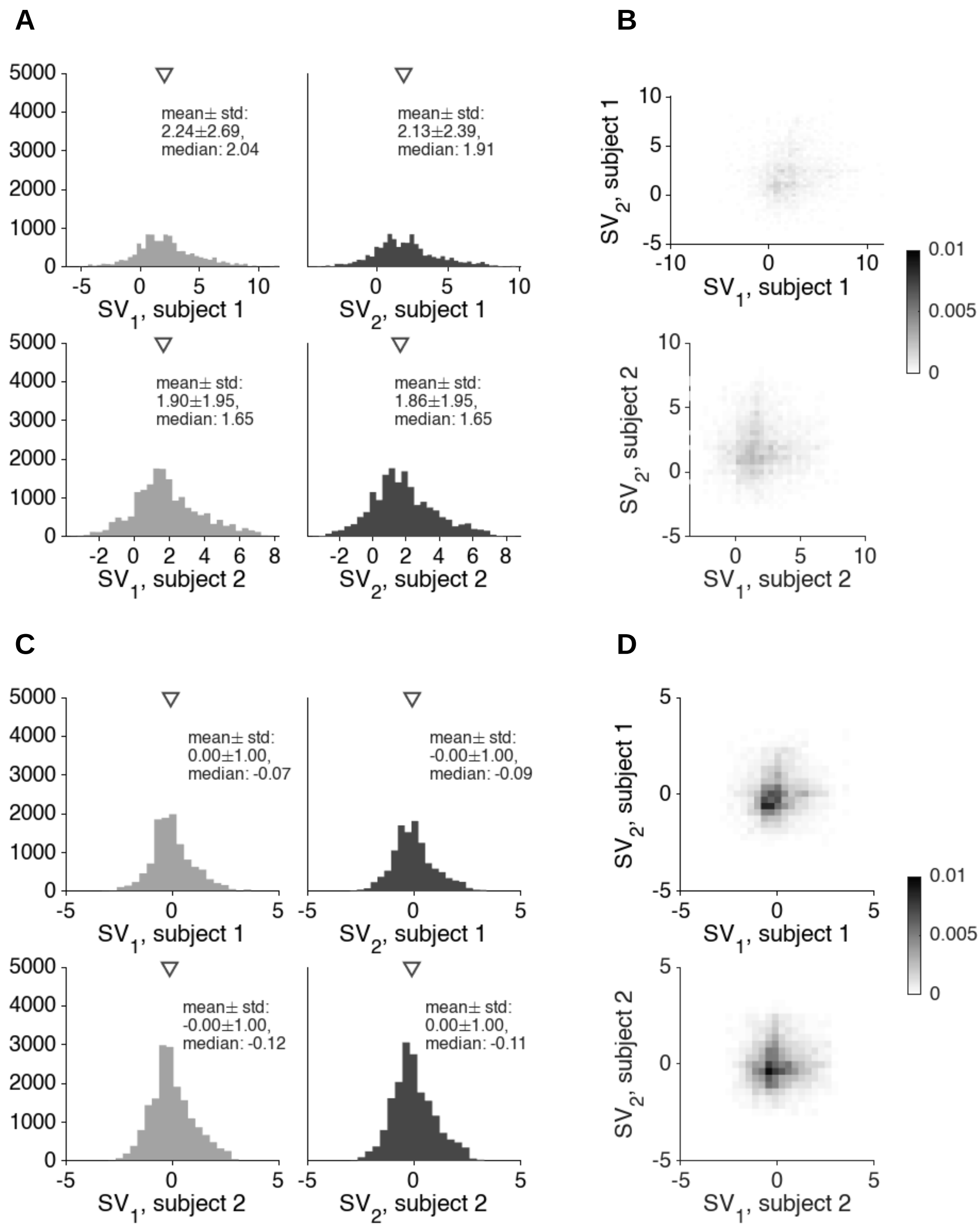
Supplementary Figure S2: Choice decoding accuracy of different alternative models. Each dot reports the choice prediction accuracy for different models, sessions, and for $\mathrm{k} = 4$ cross-validation folds in subject 1 (left) and subject 2 (right). Accuracy (mean $\pm$ sem): m_0: $60.74 \pm 0.23\%$ ($59.9 \pm 0.36\%$ for subject 1, $61.29 \pm 0.3\%$ for subject 2), m_1: $75.50 \pm 0.28\%$ ($77.58 \pm 0.5\%$ subject 1, $74.12 \pm 0.31\%$ subject 2), m_2: $82.87 \pm 0.25\%$ ($83.87 \pm 0.46\%$ subject 1, $82.21 \pm 0.28\%$ subject 2) and m_3: $82.66 \pm 0.25\%$ ($83.51 \pm 0.46\%$ subject 1, $82.09 \pm 0.29\%$ subject 2). This assessment suggests m_2 as the best model.



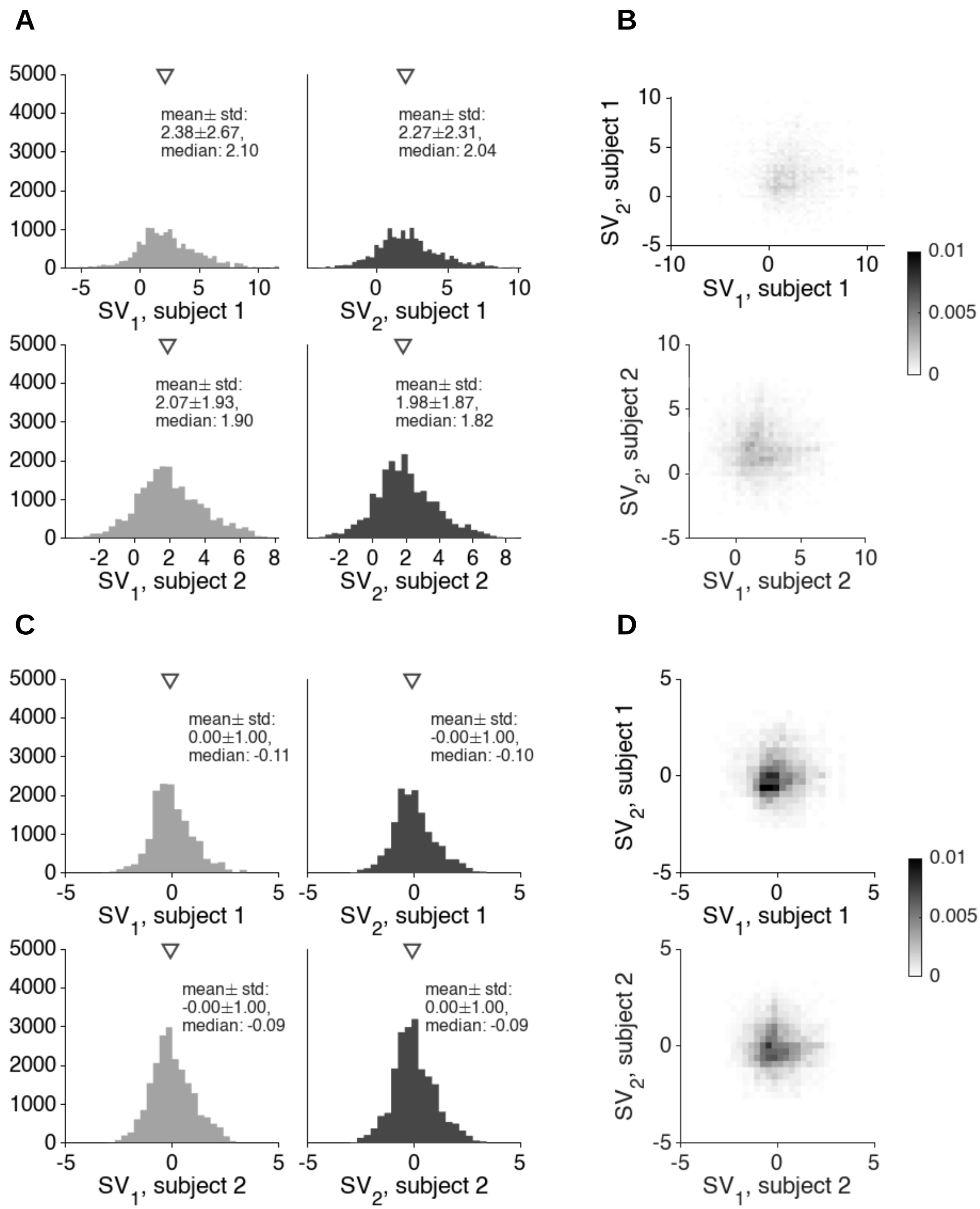
Supp. Figure S3. Distributions of SVs for Easy task configurations. A. SV_1 (left), SV_2 (right) distributions for subject 1 (top) and subject 2 (bottom). **B.** Joint distribution of SV_1, SV_2. **C-D.** Same as A-B but for z-scored data.



Supp. Figure S4. Distributions of SVs for Hard task configurations. A. SV_1 (left), SV_2 (right) distributions for subject 1 (top) and subject 2 (bottom). **B.** Joint distribution of SV_1, SV_2. **C-D.** Same as A-B but for z-scored data.



Supp. Figure S5. Distributions of SVs for Low accumulated reward task configurations. A. SV_1 (left), SV_2 (right) distributions for subject 1 (top) and subject 2 (bottom). **B.** Joint distribution of SV_1, SV_2. **C-D.** Same as A-B but for z-scored data.



Supp. Figure S6. Distributions of SVs for High accumulated reward task configurations. A. SV_1 (left), SV_2 (right) distributions for subject 1 (top) and subject 2 (bottom). **B.** Joint distribution of SV_1, SV_2. **C-D.** Same as A-B but for z-scored data.

Supplementary Methods

Linear regression of two variables

Consider $(x_1, x_2) \in \mathbb{R}^2: y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$, where $\beta_0, \beta_1, \beta_2 \in \mathbb{R}$ are unknown.

From empirical observations, we have n samples $x_{1,i}, x_{2,i}, y_i$, $i = 1, \dots, n$, and we may estimate $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$, such that $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i}, i = 1, \dots, n$.

By using Ordinary Least Squares (OLS), we derive the following (unbiased, consistent) estimators:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2,$$

$$\hat{\beta}_1 = \frac{\left(\sum_{i=1}^n (y_i - \bar{y}) x_{1,i}\right)\left(\sum_{i=1}^n (x_{2,i} - \bar{x}_2) x_{2,i}\right) - \left(\sum_{i=1}^n (y_i - \bar{y}) x_{2,i}\right)\left(\sum_{i=1}^n (x_{2,i} - \bar{x}_2) x_{1,i}\right)}{\left(\sum_{i=1}^n (x_{1,i} - \bar{x}_1) x_{1,i}\right)\left(\sum_{i=1}^n (x_{2,i} - \bar{x}_2) x_{2,i}\right) - \left(\sum_{i=1}^n (x_{1,i} - \bar{x}_1) x_{2,i}\right)\left(\sum_{i=1}^n (x_{2,i} - \bar{x}_2) x_{1,i}\right)},$$

$$\hat{\beta}_2 = \frac{\left(\sum_{i=1}^n (y_i - \bar{y}) x_{2,i}\right)\left(\sum_{i=1}^n (x_{1,i} - \bar{x}_1) x_{1,i}\right) - \left(\sum_{i=1}^n (y_i - \bar{y}) x_{1,i}\right)\left(\sum_{i=1}^n (x_{1,i} - \bar{x}_1) x_{2,i}\right)}{\left(\sum_{i=1}^n (x_{1,i} - \bar{x}_1) x_{1,i}\right)\left(\sum_{i=1}^n (x_{2,i} - \bar{x}_2) x_{2,i}\right) - \left(\sum_{i=1}^n (x_{1,i} - \bar{x}_1) x_{2,i}\right)\left(\sum_{i=1}^n (x_{2,i} - \bar{x}_2) x_{1,i}\right)},$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i$ is the sample mean of $y_i, i = 1, \dots, n$,

and $\bar{x}_1 = \frac{1}{n}\sum_{i=1}^n x_{1,i}$, $\bar{x}_2 = \frac{1}{n}\sum_{i=1}^n x_{2,i}$ are the sample means of $x_{1,i}, x_{2,i}, i = 1, \dots, n$.

The quality of the fit and the significance assessment are measured via the total sum of squares (SST), sum of squares due to regression (SSR), sum of squares due to error (SSE):

$SST = \sum_{i=1}^n (y_i - \bar{y})^2, SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2, SST = SSR + SSE.$

The significance assessment is made by testing the hypothesis that the slope terms $\hat{\beta}_j, j = 1,2$ are significantly different from zero. This is achieved by comparing the empirical $F_j{}^*$ value:

$$F_j^* = \frac{\left(SSE_j^R - SSE^F\right)/\left((n-2) - (n-3)\right)}{(SSE^F)/(n-3)}, \qquad j = 1,2,$$

to a F-distribution with $(n-2) - (n-3) = 1$ degree of freedom at the numerator, and $(n-3)$ degrees of freedom at the denominator. In this formulation, SSE^F is the SSE of the full model, that is $SSE^F = \sum_{i=1}^n \left|y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i}\right)\right|^2$, and SSE_j^R is the SSE of the reduced model for each variable x_1, x_2: $SSE_j^R = \sum_{i=1}^n \left|y_i - \left(\hat{\beta}_0 + \sum_{k=1,2:k\neq j} \hat{\beta}_k \, x_{k,i}\right)\right|^2$ for $\hat{\beta}_j, j = 1,2$.

Linear regression of two z-scored variables

Let us consider the z-scored version of x_1, x_2. Define: $z_{1,i} = \frac{x_{1,i} - \bar{x}_1}{\hat{\sigma}_{x_1}}$ and $z_{2,i} = \frac{x_{2,i} - \bar{x}_2}{\hat{\sigma}_{x_2}}$, $i = 1, \dots, n$,
where $\bar{x}_1 = \frac{1}{n}\sum_{i=1}^n x_{1,i}$ and $\bar{x}_2 = \frac{1}{n}\sum_{i=1}^n x_{2,i}$ are the sample means of x_1, x_2,
and $\hat{\sigma}_{x_1} = \frac{1}{n-1}\sum_{i=1}^n (x_{1,i} - \bar{x}_1)^2$ and $\hat{\sigma}_{x_2} = \frac{1}{n-1}\sum_{i=1}^n (x_{2,i} - \bar{x}_2)^2$ are the sample variances of x_1, x_2.

From the above, $\bar{z}_1, \bar{z}_2 = 0$, $\hat{\sigma}_{z_1} = \hat{\sigma}_{z_2} = 1$, following the definition of z-scored variables.

The previous linear regression can be applied to z-scored variables, with minor changes to the initial definition: $\hat{y}_i = \hat{\beta}'_0 + \hat{\beta}'_1 z_{1,i} + \hat{\beta}'_2 z_{2,i} = \hat{\beta}'_0 + \hat{\beta}'_1 \frac{x_{1,i} - \bar{x}_1}{\hat{\sigma}_{x_1}} + \hat{\beta}'_2 \frac{x_{2,i} - \bar{x}_2}{\hat{\sigma}_{x_2}}$.

Recalling the initial model $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i}$, we can find a direct correspondence between the intercept $\hat{\beta}_0 = \hat{\beta}'_0 - \hat{\beta}'_1 \frac{\bar{x}_1}{\hat{\sigma}_{x_1}} - \hat{\beta}'_2 \frac{\bar{x}_1}{\hat{\sigma}_{x_1}}$, and slope terms $\hat{\beta}_1 = \frac{\hat{\beta}'_1}{\hat{\sigma}_{x_1}}$, and $\hat{\beta}_2 = \frac{\hat{\beta}'_2}{\hat{\sigma}_{x_2}}$. Note that from OLS, $\hat{\beta}'_0 = \bar{y} - \hat{\beta}'_1 \bar{z}_1 - \hat{\beta}'_2 \bar{z}_2 = \bar{y}$, thus it does not depend on $\hat{\beta}'_1, \hat{\beta}'_2$.

When applying significance assessment to the initial and to the z-scored variables, we find that $F^*(z_1, z_2) = F^*(x_1, x_2)$, since the two terms SSE^F and SSE_j^R, for $j = 1,2$, coincide.

$$
\begin{aligned}
SSE^F(z_1, z_2) &= \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}'_0 + \hat{\beta}'_1 z_{1,i} + \hat{\beta}'_2 z_{2,i}\right) \right|^2 \\
&= \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}_0 + \hat{\beta}'_1 \frac{\bar{x}_1}{\hat{\sigma}_{x_1}} + \hat{\beta}'_2 \frac{\bar{x}_2}{\hat{\sigma}_{x_2}} + \hat{\beta}'_1 \frac{x_{1,i} - \bar{x}_1}{\hat{\sigma}_{x_1}} + \hat{\beta}'_2 \frac{x_{2,i} - \bar{x}_2}{\hat{\sigma}_{x_2}}\right) \right|^2 \\
&= \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}_0 + \hat{\beta}'_1 \frac{x_{1,i}}{\hat{\sigma}_{x_1}} + \hat{\beta}'_2 \frac{x_{2,i}}{\hat{\sigma}_{x_2}}\right) \right|^2 \\
&= \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i}\right) \right|^2 = SSE^F(x_1, x_2),
\end{aligned}
$$

In this case, for the SSE_1^R, we compare the full model with the reduced models $\hat{y}_i = \hat{\beta}'_0 + \hat{\beta}'_2 z_{2,i}$ and $\hat{y}_i = \hat{\beta}'_0 + \hat{\beta}'_2 x_{2,i}$, where the variables z_1, x_1 are omitted. This time we have $\hat{\beta}'_0 = \hat{\beta}_0 + \hat{\beta}'_2 \frac{\bar{x}_2}{\hat{\sigma}_{x_2}}$.

$$
\begin{aligned}
SSE_1^R(z_1, z_2) &= \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}'_0 + \hat{\beta}'_2 z_{2,i}\right) \right|^2 \\
&= \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}_0 + \hat{\beta}'_2 \frac{\bar{x}_2}{\hat{\sigma}_{x_2}} + \hat{\beta}'_2 \frac{x_{2,i} - \bar{x}_2}{\hat{\sigma}_{x_2}}\right) \right|^2 \\
&= \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}_0 + \hat{\beta}'_2 \frac{x_{2,i}}{\hat{\sigma}_{x_2}}\right) \right|^2 = \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}_0 + \hat{\beta}_2 x_{2,i}\right) \right|^2 = SSE_1^R(x_1, x_2),
\end{aligned}
$$

Similarly, for SSE_2^R we compare the full model with reduced models omitting the variables z_2, x_2:

$$
SSE_2^R(z_1, z_2) = \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}'_0 + \hat{\beta}'_1 z_{1,i}\right) \right|^2 = \cdots = \sum_{i=1}^{n} \left| y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_{1,i}\right) \right|^2 = SSE_2^R(x_1, x_2).
$$

The reduced model SSE for z-scored variables coincides with the SSE for the initial data.

Furthermore, one may note that F^* mainly depends on SSE, and that the total sum of squares $SST = SSR + SSE = \sum_{i=1}^{n}(y_i - \bar{y})^2$ does not depend on the regressed variables (x_1, x_2) or (z_1, z_2), thus $F^*(z_1, z_2) = F^*(x_1, x_2)$, implies that $SSR(z_1, z_2) = SSR(x_1, x_2)$.

$$
\begin{aligned}
SSR(z_1, z_2) = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 &= \sum_{i=1}^{n}\left(\hat{\beta}_0' + \hat{\beta}_1' z_{1,i} + \hat{\beta}_2' z_{2,i} - \bar{y}\right)^2 \\
&= \sum_{i=1}^{n}\left(\hat{\beta}_0' + \hat{\beta}_1' \frac{x_{1,i} - \bar{x}_1}{\hat{\sigma}_{x_1}} + \hat{\beta}_2' \frac{x_{2,i} - \bar{x}_2}{\hat{\sigma}_{x_2}} - \bar{y}\right)^2 \\
&= \sum_{i=1}^{n}\left(\hat{\beta}_1' \frac{x_{1,i} - \bar{x}_1}{\hat{\sigma}_{x_1}} + \hat{\beta}_2' \frac{x_{2,i} - \bar{x}_2}{\hat{\sigma}_{x_2}}\right)^2,
\end{aligned}
$$

$$
\begin{aligned}
SSR(x_1, x_2) = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 &= \sum_{i=1}^{n}\left(\hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i} - \bar{y}\right)^2 \\
&= \sum_{i=1}^{n}\left(\hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2\right)^2,
\end{aligned}
$$

$$SSR(z_1, z_2) = SSR(x_1, x_2).$$

This latter result implies that also the coefficient of determination $R^2 = SSR/SST$ coincides in the two models, i.e., $R^2(z_1, z_2) = R^2(x_1, x_2)$.