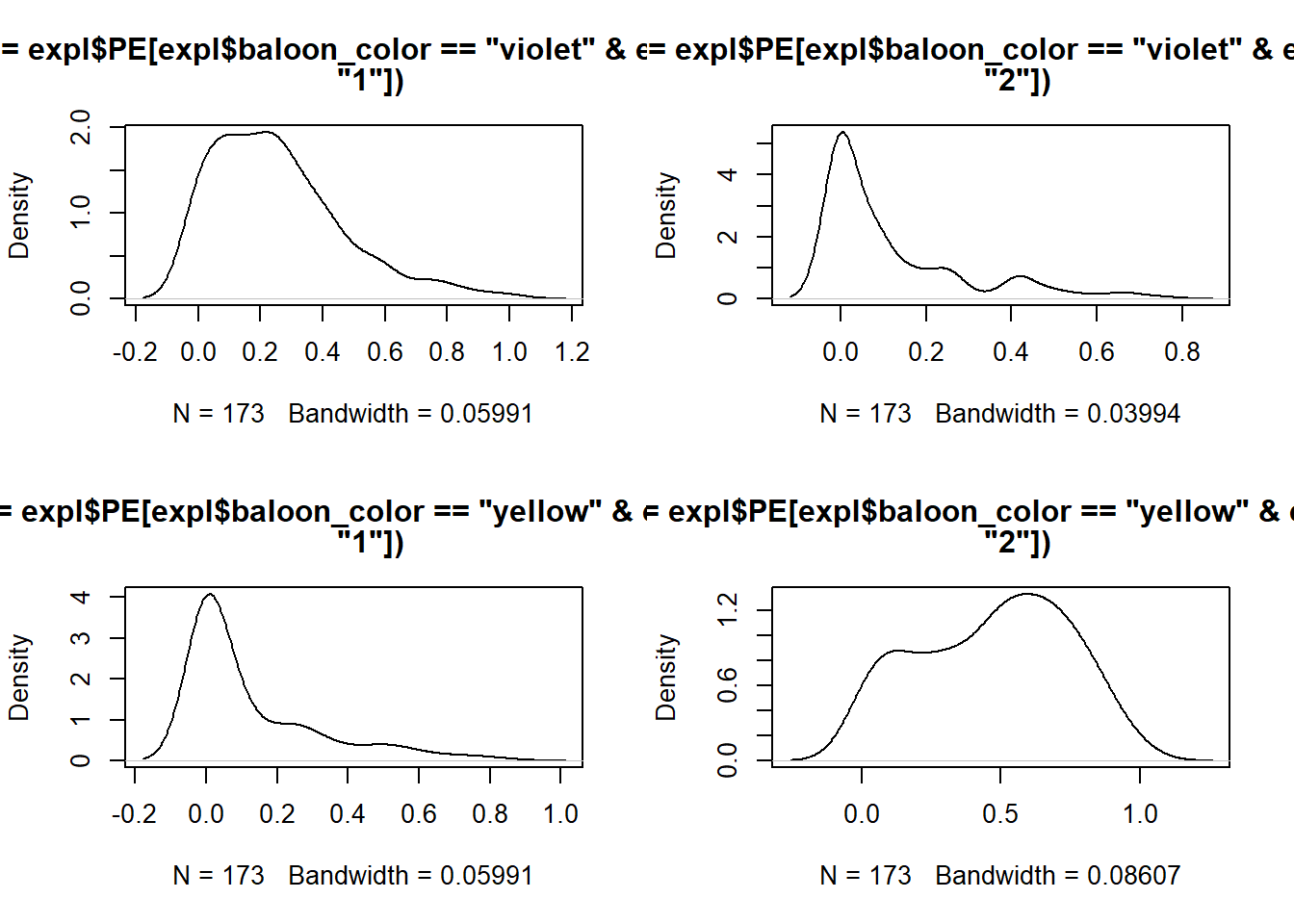
**Supporting Information**

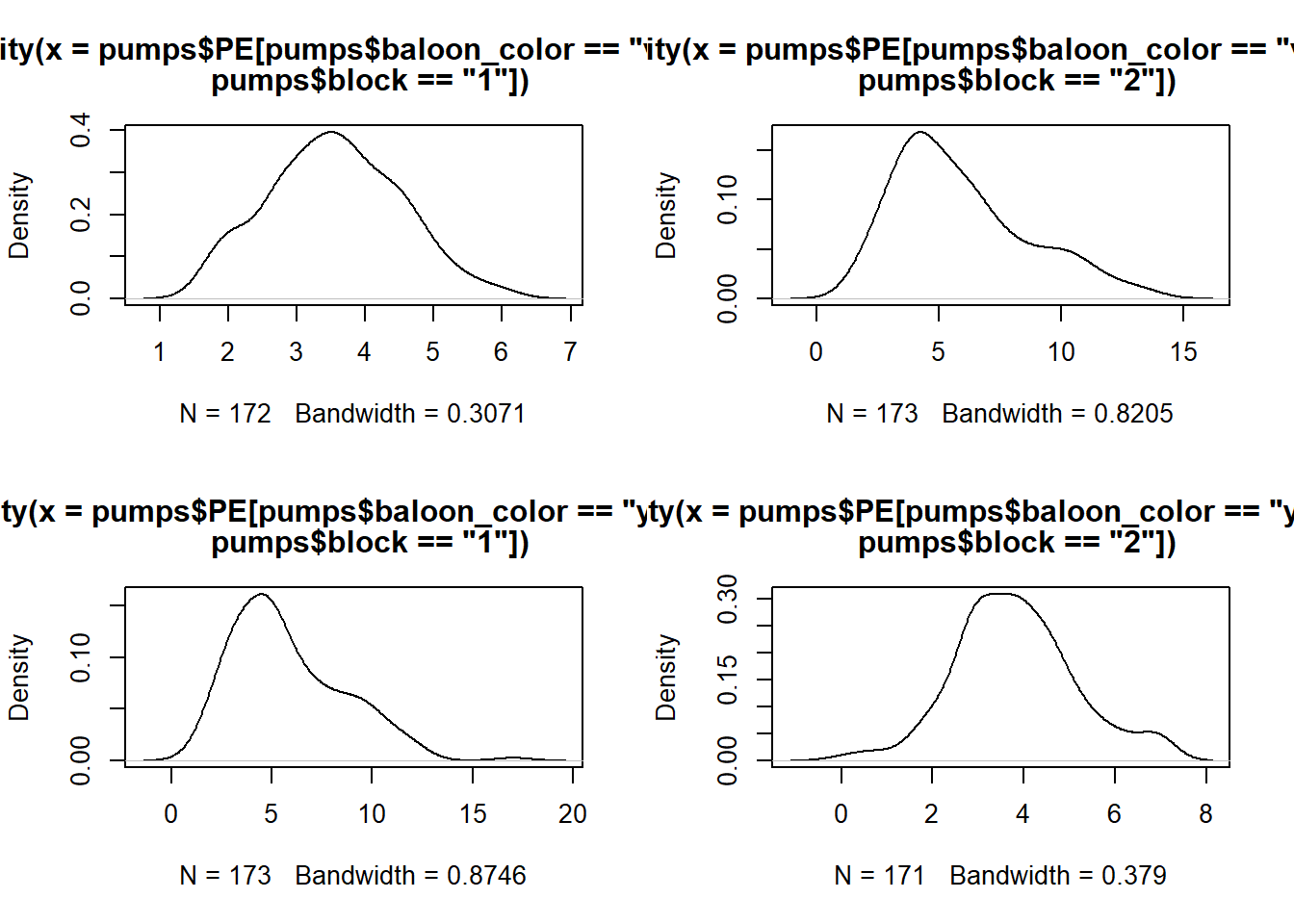
The data and analysis script to reproduce these model selections will be made openly available in Open Science Framework (OSF) and can be accessed at: <https://osf.io/shuq2/?view_only=6676ca3e2a6e426c8a61839d24cc37f1>

# **S1. Adjusted pumps and explosions distribution**

Here we report the density graph of the pumps (only in successful trials) and of explosions distributed in the four conditions: violet low advantageous (block 1 ‘YELLOWviolet’, upper panel left), violet high advantageous (block 2 ‘yellowVIOLET’, upper panel right), yellow high advantageous (block 1 ‘YELLOWviolet’, lower panel left) and yellow low advantageous (block 2 ‘yellowVIOLET’, lower panel right).

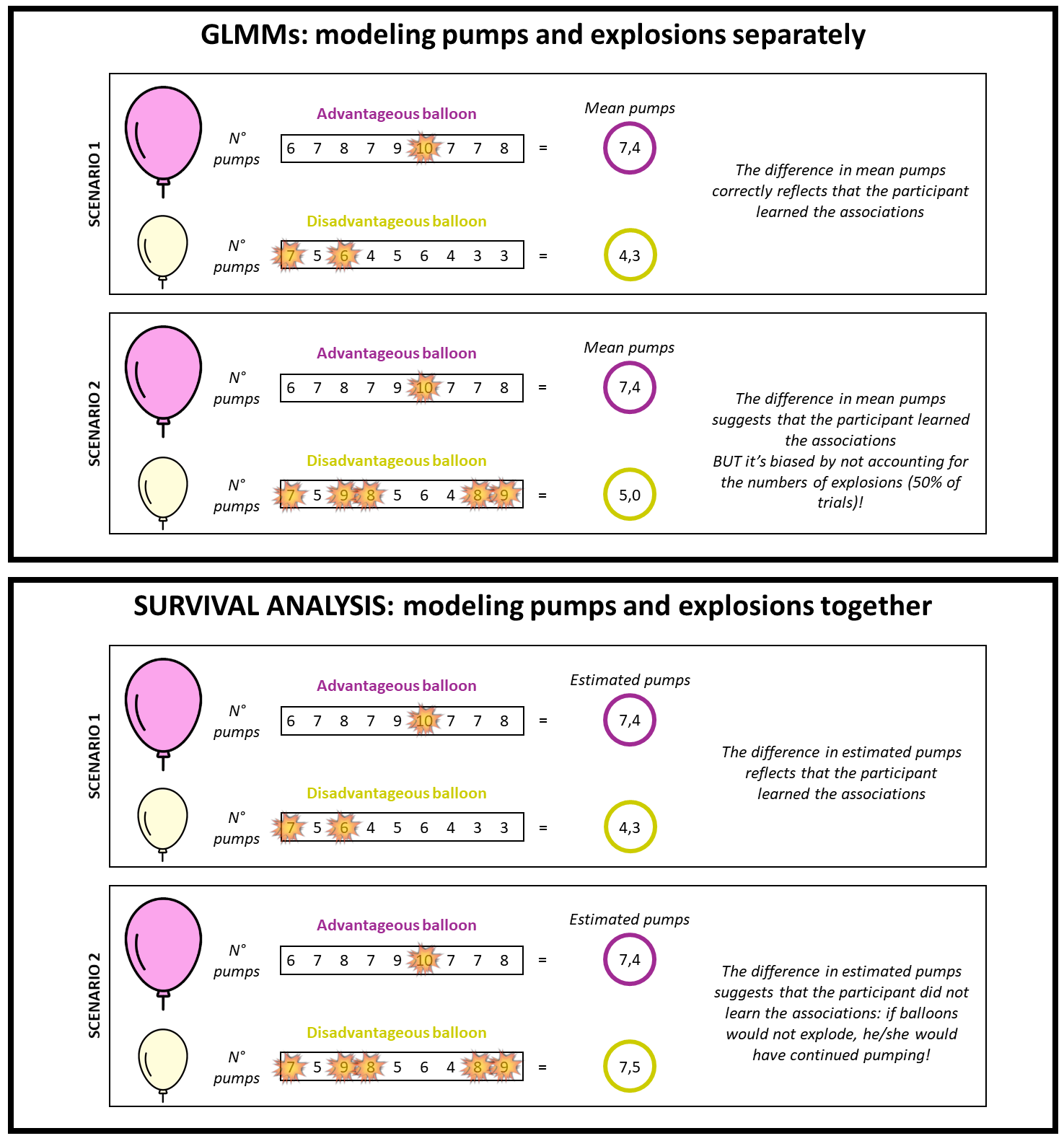


**Figure S1.1. Explosions.** *Upper panel.* Explosion distribution for the violet balloons in block 1 (left) and block 2 (right). *Bottom panel.* Explosion distribution for the yellow balloons in block 1 (left) and block 2 (right).



**Figure S1.2. Pumps in successful trials.** *Upper panel.* Pumps distribution for the violet balloons in block 1 (left) and block 2 (right). *Bottom panel.* Pumps distribution for the yellow balloons in block 1 (left) and block 2 (right).

# **S2. Modelling pumps and explosions**

****

**Figure S2.1. Modelling pumps and explosions: an example.** Here we report an example of the implications of modelling pumps and explosions separately using generalized linear mixed effects models (GLMMs) vs together using the survival analysis.

# **S3. Model selection results**

**S3.1 Survival analysis.**

All models were fitted using an Exponential family distribution with a log link function, weakly informative prior for the regression coefficient of the interaction term (gaussian(0,1)) and brms default priors for the other parameters (i.e., correlations of group-level ('random') effects: lkj\_corr\_cholesky(1); standard deviations of group-level ('random') effects: student\_t(3, 0, 2.5)). We ran 4 chains, with 4000 samples each and with the first 2000 discarded as burn-in.

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **ΔELPD** | **Δ*SE*** | **Right side of the formula** |
| **mod3** | 0.0 | 0.0 | 1 + color \* block + age + (1 + color \* block |subj) |
| **mod4** | -0.1 | 0.9 | 1 + color \* block \* age + (1 + color \* block |subj) |
| **mod\_null** | -116.4 | 9.9 | 1 + (1|subj) |

**Table S3.1. Survival analysis.** We see from [Table](https://pubs.asha.org/doi/full/10.1044/2018_JSLHR-S-18-0006?casa_token=RFKOI_LC0m4AAAAA%3AKuIp4mUBW3YxN9Wd7_1N7mSk3HdkO_6mCC3dO_XG_1g_7gO6d-_IpDJjidLqoSrx815K0zWbsRiaZA#T7) 1 that mod3 is the best performing model (see ELPD and SE values).

**S3.2 Questionnaires.**

The models were fitted using a Gaussian family distribution, weakly informative priors for the regression coefficients of cluster (i.e., student\_t(3,0,3)) and for the intercepts (i.e., all with a lower bound on zero and means coming from gaussian distributions: ADHD Index: gaussian(6,3); DSM disattention: gaussian(5,4); DSM impulsivity and hyperactivity: gaussian(6,5); CGI Restlessness/Hyperactivity: gaussian(4,3)) and brms default priors for all the other parameters (i.e., residual standard deviation: ADHD Index: student\_t(3, 0, 5.9); DSM disattention: student\_t(3, 0, 3.7); DSM impulsivity and hyperactivity: student\_t(3, 0, 4.4); CGI Restlessness/Hyperactivity: student\_t(3, 0, 3)). We ran 4 chains, with 6000 samples each and with the first 2000 discarded as burn-in.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **LOOIC** | ***SE LOOIC*** | **ΔELPD** | **Δ*SE*** | **Right side of the formula** |
| **mod1** | 626.726 | 18.511 | 0.000 | 0.000 | 1 + cluster |
| **mod3** | 628.575 | 18.366 | -0.924 | 0.233 | 1 + cluster + age |
| **mod\_null** | 629.628 | 19.587 | -1.451 | 2.098 | 1 |
| **mod2** | 631.355 | 19.402 | -2.315 | 2.045 | 1 + age |

**Table S3.2.1. ADHD index.** For the ADHD index composite scale, mod1 is performing better than the other models, as it has the lower LOOIC and ELPD.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **LOOIC** | ***SE LOOIC*** | **ΔELPD** | **Δ*SE*** | **Right side of the formula** |
| **mod\_null** | 552.917 | 15.325 | 0.000 | 0.000 | 1 |
| **mod1** | 553.366 | 15.228 | -0.224 | 1.293 | 1 + cluster |
| **mod2** | 553.748 | 15.442 | -0.415 | 1.960 | 1 + age |
| **mod3** | 554.420 | 15.372 | -0.751 | 1.677 | 1 + cluster + age |

**Table S3.2.2. DSM disattention.** For the DSM Disattention composite scale, mod\_null is performing better than the other models, as it has the lower LOOIC and ELPD.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **LOOIC** | ***SE LOOIC*** | **ΔELPD** | **Δ*SE*** | **Right side of the formula** |
| **mod1** | 591.021 | 17.664 | 0.000 | 0.000 | 1 + cluster |
| **mod\_null** | 591.377 | 18.696 | -0.178 | 1.514 | 1 |
| **mod3** | 591.523 | 17.307 | -0.251 | 0.994 | 1 + cluster + age |
| **mod2** | 591.787 | 18.217 | -0.383 | 1.588 | 1 + age |

**Table S3.2.3. DSM hyperactivity and impulsivity.** For the DSM Hyperactivity and Impulsivity composite scale, mod1 is performing better than the other models, as it has the lower LOOIC and ELPD.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **LOOIC** | ***SE LOOIC*** | **ΔELPD** | **Δ*SE*** | **Right side of the formula** |
| **mod1** | 524.453 | 16.940 | -0.000 | 0.000 | 1 + cluster |
| **mod\_null** | 525.562 | 17.975 | -0.554 | 1.654 | 1 |
| **mod3** | 526.659 | 16.831 | -1.103 | 0.150 | 1 + cluster + age |
| **mod2** | 527.363 | 17.812 | -1.455 | 1.608 | 1 + age |

**Table S3.2.4. CGI restlessness/hyperactivity.** For the CGI restlessness/hyperactivity composite scale, mod1 is performing better than the other models, as it has the lower LOOIC and ELPD.

# **S4. Number of clusters and cluster stability**

We used the NbClust package (Charrad et al., 2014) to identify the best clustering solutions on 100 randomly selected subsets (50% of the whole dataset). Overall, in 52% of the iterations 2 was the best solution, in 47% 3 was the best solution and in 1% of iterations 6 was the best solution. Therefore, although 3 might be a plausible solution too, 2 was a more stable and parsimonious solution. To check the stability of cluster assignment, we run the k-means clustering on 100 randomly selected subsets (50% of the whole dataset). At each iteration, we compared the current cluster assignment (i.e., either 1 or 2) with the original cluster assignment (reported in the Manuscript). Therefore, at each iteration and for each subject, we could estimate the level of accordance between the current and the original cluster assignment and retrieve a mean measure of accordance for the whole sample. Importantly: labels (1 vs 2) are random, so that cluster 1 in one iteration can correspond to cluster 2 in another iteration. Therefore, a very low mean accordance (e.g., < .20) or a very high accordance (e.g., > .80) can be considered as suggesting high coherence in label assignment (if accordance is < .20 it suggests that participants have been assigned to the same cluster as the original, but labels are inverted). Overall, in 96% of iterations, coherence of cluster assignment was over 80%. Please find the script and results at the following link: <https://osf.io/shuq2/?view_only=6676ca3e2a6e426c8a61839d24cc37f1> 🡪 “Analysis” 🡪 “Scripts” 🡪 ”Scripts and analyses for reviewers” 🡪 “kmeans\_clustering\_checks\_restReviewers\_june24.html”.

# **S5. Experimental design and time-on-task effects**

**S5.1 Experimental design**

It is important to note that analyzing the interaction between color and advantageousness versus color and block is effectively equivalent, as both approaches examine the data from different perspectives. In our experimental design, the variables of color, block, and advantageousness are interdependent, as illustrated in the accompanying Figure:

****

**Figure S5.1.1.** Experimental design. The figure describes the experimental design. From the picture it is clear that some of the levels are missing (i.e., violet advantageous block 1, violet disadvantageous block 2, yellow advantageous block 2, yellow disadvantageous block 1). Therefore, violet advantageous corresponds to violet in block 2, violet disadvantageous corresponds to violet in block 1 and so forth.

In other words, testing e.g., the difference between violet disadvantageousness and violet advantageously is the same as testing violet in Block 1 vs violet in Block 2. However, when examining interaction effects, we observe an interaction for color X advantageousness and for block X advantageousness but not for color X block, as illustrated in the following Figure:

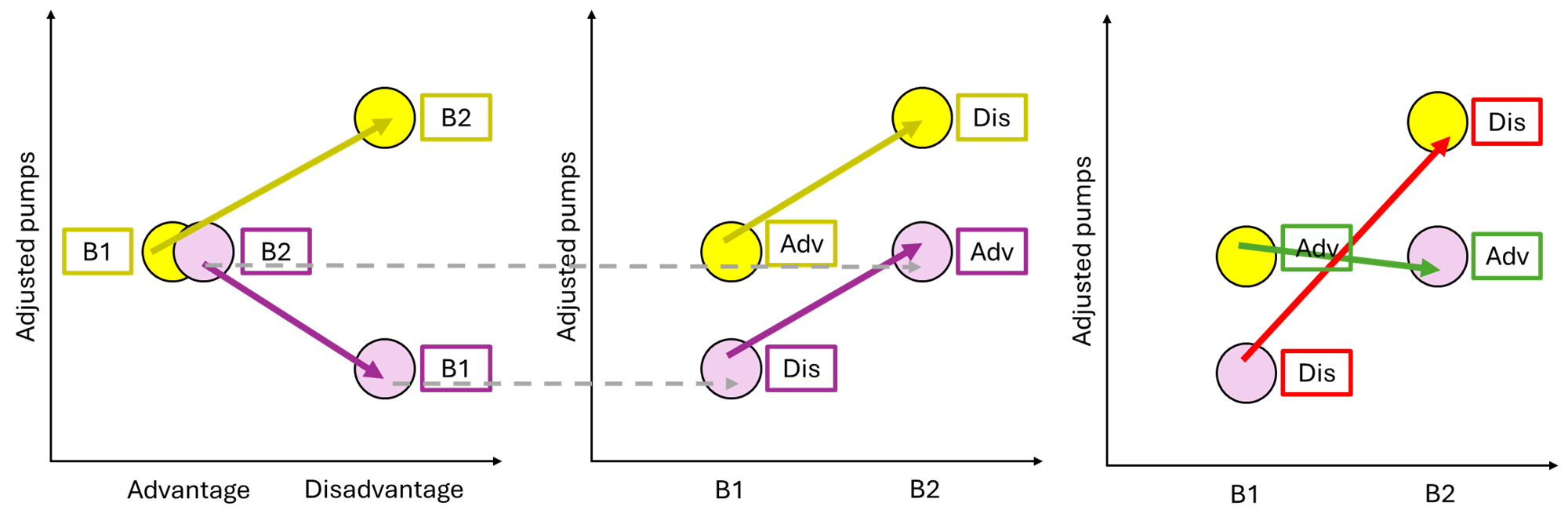


Figure S5.1.2. Graphical comparison of color X advantageousness versus color X block and block X advantageousness. Left figure: adjusted pumps are displayed on the y-axis, advantageousness on the x-axis. The picture shows a decrease in pumps for violet balloons and an increase in pumps for yellow balloons when going from the advantage to the disadvantage condition. Yellow advantage and violet disadvantage correspond to block 1 (B1) and yellow disadvantage and violet advantage correspond to block 2 (B2). Middle figure: adjusted pumps are displayed on the y-axis, block on the x-axis. The picture shows an overall increase in pumps for both violet and yellow balloons when going from block 1 to block 2. Yellow B1 and violet B2 correspond to advantage (Adv) and yellow B2 and violet B1 correspond to disadvantage (Dis). Right figure: adjusted pumps are displayed on the y-axis, block on the x-axis. The picture shows an overall increase in pumps for disadvantageous balloons (red arrow) when going from block 1 to block 2, and no difference for advantageous balloons (green arrow). Yellow B1 and violet B2 correspond to advantage (Adv) and yellow B2 and violet B1 correspond to disadvantage (Dis).

**S5.2 Considering advantageousness instead**  **of block**

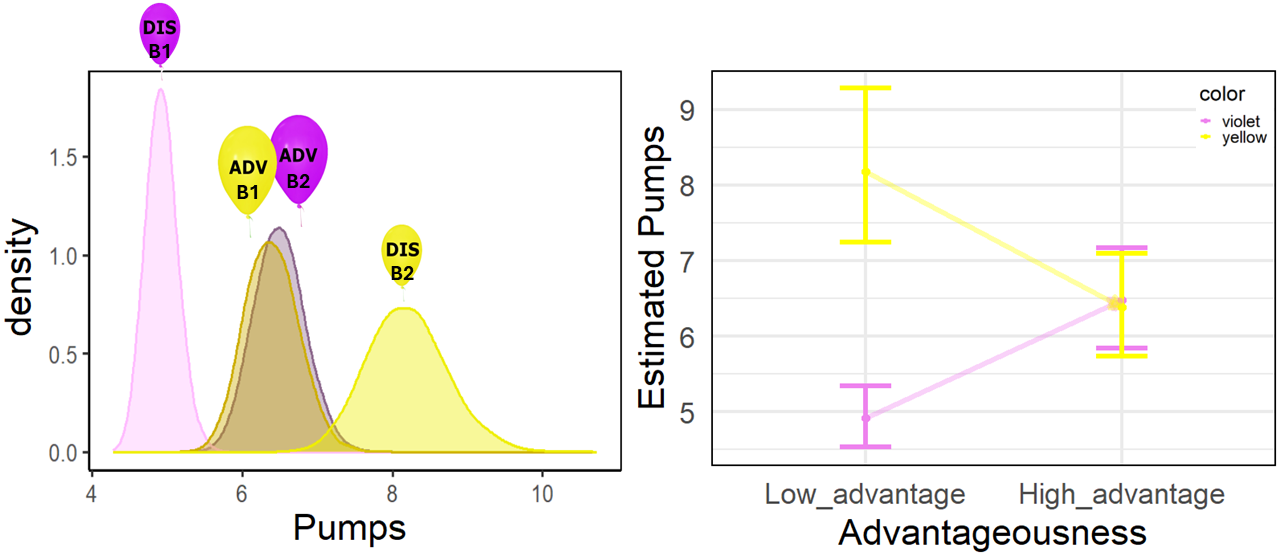
Here we report the results of the color X advantageousness model:

* dependent variable = “pumps | cens(explosions)”, number of pumps conditioned by the number of explosions
* independent variables = “color” (factor with 2 levels: violet, yellow) X “advantageousness ” (factor with 2 levels: low , high ), and “age” (numeric scaled) as covariate
* random part = “1 + color X advantageousness | subject”, both group and individual parameters for each interaction levels

|  |  |  |  |
| --- | --- | --- | --- |
|  | ꞵ | PI | MPE |
| color (violet - yellow) | -0.25 | [-0.38 , -0.15 ] | >99% |
| advantageousness (low - high ) | -0.01 | [-0.12 , 0.10 ] | 68.08 % |
| age | -0.02 | [-0.19 , 0.14 ] | 73.39 % |
| color X advantageousness | -0.52 | [-0.77 , -0.21 ] | >100 % |

**Table S5.2.** For each main effect and for the interaction term, in the table we reported the ꞵ parameter, the full posterior interval (PI) and the Maximum Probability of an Effect (MPE).

We found a plausible main effect of color (overall, higher number of adjusted pumps for the yellow compared to the violet balloon) but not of advantageousness (overall, the number of pumps in advantageous compared to disadvantageous balloons do not differ ). However, we found a plausible effect of the color X advantageousness interaction (overall, higher number of pumps for yellow advantageous compared to disadvantageous balloons, and the opposite patterns for violet balloons). We did not find a plausible main effect of age (see the Figures below).



**Figure S5.2.** *Left figure*: number of pumps estimated by the survival analysis (x-axis) and posterior distribution densities (y-axis) for the four-levels of the color X advantageousness interaction: yellow high advantage (yellow with large yellow balloon on top), yellow low advantage (yellow with small yellow balloon on top), violet high advantage (violet with large violet balloon on top), violet low advantage (violet with small violet balloon on top). Labels B1 and B2 stand for block 1 and block 2, respectively. *Right figure*: Mean and standard deviation of the estimated pumps by the survival analysis (y-axis) and advantageousness (x-axis) for the two balloon colors (violet and yellow). The arrows show an increase in estimated pumps from low to high advantage for the violet balloons, while showing the opposite pattern for yellow baloons .

For the sake of completeness, we also report here the block X advantageousness interaction.

* dependent variable = “pumps | cens(explosions)”, number of pumps conditioned by the number of explosions
* independent variables = “adv” (factor with 2 levels: high, low) X “block” (factor with 2 levels: 1, 2), and “age” (numeric scaled) as covariate
* random part = “1 + adv X block | subject”, both group and individual parameters for each interaction levels

|  |  |  |  |
| --- | --- | --- | --- |
|  | ꞵ | PI | MPE |
| adv (low - high) | -0.01 | [-0.13, 0.08] | 67.95% |
| block (block 1 - block 2) | -0.26 | [-0.39, -0.15] | >99% |
| age | -0.03 | [-0.20, 0.14] | 74.16% |
| adv (low - high) X block (block 1 - block 2) | -0.49 | [-0.71, -0.26] | >99% |

**Table S5.3.** For each main effect and for the interaction term, in the table we reported the ꞵ parameter, the full posterior interval (PI) and the Maximum Probability of an Effect (MPE).

Immagine che contiene testo, diagramma, Diagramma, linea

Descrizione generata automaticamenteImmagine che contiene testo, diagramma, linea, Diagramma

Descrizione generata automaticamente

**Figure S5.3.** *Left figure*: number of pumps estimated by the survival analysis (x-axis) and posterior distribution densities (y-axis) for the four-levels of the color X block interaction: yellow high advantage (light yellow), yellow low advantage (dark yellow), violet high advantage (dark violet), violet low advantage (light violet). Labels B1 and B2 stand for block 1 and block 2, respectively. *Right figure*: Mean and standard deviation of the estimated pumps by the survival analysis (y-axis) and blocks (x-axis) for advantageous and disadvantageous balloons. The arrows show an increase in estimated pumps in the disadvantageous condition from block 1 to block 2, but not for the disadvantageous condition.

In this case, as with the color X advantageousness interaction, we find an interaction suggesting that while children made a similar number of pumps in advantageous conditions across both blocks, this was not the case for disadvantageous conditions, where overall pumps increased in Block 2. If the observed increase in pumps was solely due to greater confidence over time, we would have expected a higher number of pumps in the advantageous condition of Block 2 compared to Block 1.

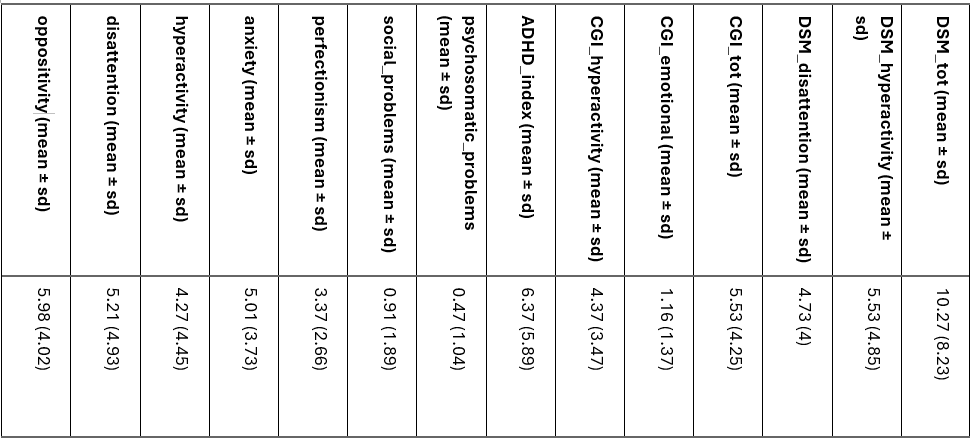
# **S5** **. BART and CPRS descriptive statistics**

Descriptive statistics of BART measures on the whole sample:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | Mean pumps ± SD | | | | % explosions (mean ± SD) | | | | Mean score ± SD (range) |
| YH | YL | VH | VL | YH | YL | VH | VL |
| 170 | 5.7 ± 2.8 | 3.8 ± 1.3 | 5.8 ± 2.7 | 3.6 ± 1.0 | 11.9% ± 18.7 | 46.3% ± 26.2 | 11.2% ± 16.6 | 24.6% ± 19.7 | 166 ± 35  (82-239) |

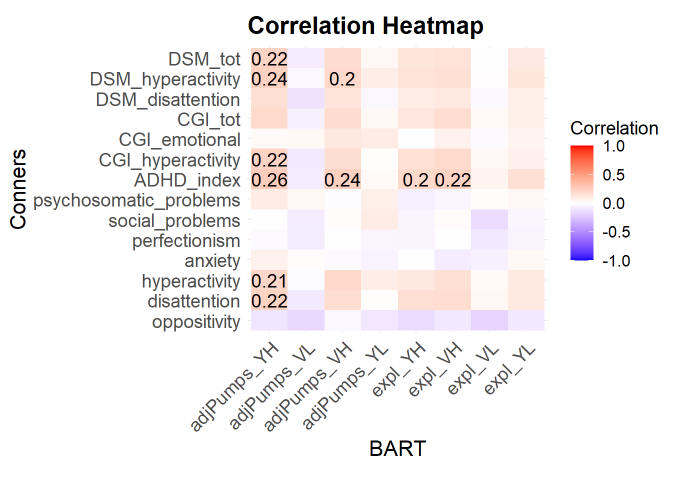
**Table S5** **.1. BART descriptive statistics.** For each condition (YH = yellow high advantage block 1, YL = yellow low advantage block 2, VH = violet high advantage block 2, VL = violet low advantage block 1), we report the mean number of pumps ± standard deviations and the %, mean ± standard deviations of explosions. Moreover, we report the total number of the sample (N) and the mean ± standard deviations and min-max range of the total score.

Descriptive statistics of CPRS scores on the whole sample (mean (sd)):



**Table S5** **.2.1. CPRS descriptive statistics.** For each subscale, we report the mean ± standard deviations across the whole sample.

Correlations between BART measures and CPRS scores (only correlations >.20 or <.20 are reported):



**Figure S5** **.2.2. Correlation between BART and CPRS.** Each row represents a different CPRS subscale, each column represents a different BART measure. Blue colors indicate negative relationships, red colors indicate positive relationships. The correlation coefficient has been specified only for significant correlations (correlation coefficient > or < 0.2).

You can find the results at the following link:

<https://osf.io/shuq2/?view_only=6676ca3e2a6e426c8a61839d24cc37f1> 🡪 “Analysis” 🡪 “Scripts” 🡪 ”Scripts and analyses for reviewers” 🡪 “demo\_correlations\_respReviewers\_june24.html”.

**S6** **. Frequentist approach**

Generalized Mixed Effects Models (GLMMs)

Pumps

* Dependent variable = pumps
* Independent variable = block X block
* Random effects = random intercept and random slopes for the interaction between block and block for each subject, allowing for individual variation in both baseline responses and the relationship between block, color , and their interaction.
* Family = Poisson

|  |  |  |  |
| --- | --- | --- | --- |
|  | ꞵ | z | p-value |
| color (violet - yellow) | -0.02 | -1.56 | 0.12 |
| block (1 - 2 ) | -0.05 | -2.72 | = .007 |
| color (violet - yellow) X  block (1 - 2 ) | -0.79 | -15.40 | < .001 |

**Table S7.1.** In the table is reported the ꞵ coefficient, z value and p value for each predictor.

Post-hoc contrasts:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| contrast | | estimate | SE | df | z | p-value |
| block 1 - block 2 | violet | -0.44 | 0.03 | Inf | -14.22 | <.001 |
| yellow | -0.34 | 0.03 | Inf | 10.81 | <.001 |
| violet - yellow | block 1 | -0.42 | 0.03 | Inf | -13.50 | <.001 |
| block 2 | 0.37 | 0.03 | Inf | 13.52 | <.001 |

**Table S7.2.** Post-hoc contrasts. In the table is reported the estimate, standard error (SE), degrees of freedom (df), z-test and p-value for each contrats.

Explosions

* Dependent variable = explosions
* Independent variable = block X color
* Random effects = random intercept and random slopes for the interaction between block and color for each subject, allowing for individual variation in both baseline responses and the relationship between block, color , and their interaction.
* Family = Binomial

|  |  |  |  |
| --- | --- | --- | --- |
|  | ꞵ | z | p-value |
| color (violet - yellow) | -0.54 | -4.91 | <.001 |
| block (1 - 2 ) | -0.59 | -4.86 | <.001 |
| color (violet - yellow) X  block (1 - 2 ) | 4.74 | 17.21 | <.001 |

**Table S7.3.** In the table is reported the ꞵ coefficient, z value and p value for each predictor.

Post-hoc contrasts:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| contrast | | estimate | SE | df | z | p-value |
| block 1 - block 2 | violet | 1.78 | 0.18 | Inf | 9.87 | <.001 |
| yellow | -2.96 | 0.19 | Inf | -15.81 | <.001 |
| violet - yellow | block 1 | 1.83 | 0.19 | Inf | 9.80 | <.001 |
| block 2 e | -2.91 | 0.17 | Inf | -17.48 | <.001 |

**Table S7.4.** Post-hoc contrasts. In the table is reported the estimate, standard error (SE), degrees of freedom (df), z-test and p-value for each contrats.

You can find an in-depth discussion of the data analysis, including various applications and results supporting our findings, at the following link:

<https://osf.io/shuq2/?view_only=6676ca3e2a6e426c8a61839d24cc37f1> 🡪 “Analysis” 🡪 “Scripts” 🡪 ”Scripts and analyses for reviewers” 🡪 “brms\_survival\_respReviewers\_june24.html”.