



OPEN

## Reliance on model-based and model-free control in obesity

Lieneke K. Janssen<sup>1,2,✉</sup>, Florian P. Mahner<sup>2</sup>, Florian Schlagenhaut<sup>2,3</sup>, Lorenz Deserno<sup>1,2,4,5,6,8</sup> & Annette Horstmann<sup>1,2,7,8</sup>

Consuming more energy than is expended may reflect a failure of control over eating behaviour in obesity. Behavioural control arises from a balance between two dissociable strategies of reinforcement learning: model-free and model-based. We hypothesized that weight status relates to an imbalance in reliance on model-based and model-free control, and that it may do so in a linear or quadratic manner. To test this, 90 healthy participants in a wide BMI range [normal-weight ( $n = 31$ ), overweight ( $n = 29$ ), obese ( $n = 30$ )] performed a sequential decision-making task. The primary analysis indicated that obese participants relied less on model-based control than overweight and normal-weight participants, with no difference between overweight and normal-weight participants. In line, secondary continuous analyses revealed a negative linear, but not quadratic, relationship between BMI and model-based control. Computational modelling of choice behaviour suggested that a mixture of both strategies was shifted towards less model-based control in obese participants. Our findings suggest that obesity may indeed be related to an imbalance in behavioural control as expressed in a phenotype of less model-based control potentially resulting from enhanced reliance on model-free computations.

Obesity is the result of systematically consuming more energy than is expended. This can be seen as a failure of control over eating behaviour<sup>1–3</sup> and could result from altered processing of reward<sup>4</sup>. As a consequence, appetitive and often high-caloric foods are over-consumed despite negative consequences, such as the uncomfortable feeling of being full, feelings of regret, or long-term health risks. Such failures of behavioural control in obesity may arise from alterations in reinforcement learning<sup>5</sup>. Indeed, obesity-related impairments in reward- and punishment-based cue-conditioning have been observed in the context of both food and monetary outcomes<sup>6</sup>, as well as impairments in appetitive conditioning in the context of chocolate rewards<sup>7</sup> (but see<sup>8</sup>). Furthermore, obese participants exhibited impairments in learning from negative outcomes when money or points served as an incentive<sup>6,9,10</sup>. These studies have focused on forms of learning that mostly resemble retrospective model-free ‘trial-and-error’ reinforcement learning. However, behavioural control arises from a balance between model-based and model-free control<sup>11,12</sup>. Model-based control relies on an internal model of the environment to enable forward planning. As a result, this system is flexible (but cognitively costly), allowing us to be goal-directed even when the environment changes, e.g. abrupt change in the current outcome value, changes. In contrast, the model-free system is cognitively inexpensive and fast (but inflexible) and is thought to underlie habitual control. To better understand this balance in obesity, the current study investigates relative reliance on model-based and model-free control of choice behaviour.

Indirect evidence links obesity to reduced model-based, or rather, goal-directed control. Previous outcome devaluation studies tapping into goal-directed and habitual control of food choices in obesity have shown a negative correlation between goal-directed control and degree of obesity in humans<sup>13,14</sup>. That is, the higher the BMI, the less participants adjusted their food choices after devaluation of one of the two choices. Behavioural adjustment after outcome devaluation of non-food rewards related positively to model-based, but not model-free control, in healthy human participants performing a two-step decision-making task<sup>15–17</sup> (but see<sup>18</sup>). Alterations in model-based vs. model-free control have been associated with behavioural inflexibility as observed in clinical

<sup>1</sup>Integrated Research and Treatment Center Adiposity Diseases, Leipzig University Medical Center, Leipzig, Germany. <sup>2</sup>Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. <sup>3</sup>Department of Psychiatry and Psychotherapy, Charité-Universitätsmedizin Berlin, Campus Charité Mitte, Berlin, Germany. <sup>4</sup>Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London, UK. <sup>5</sup>The Wellcome Centre for Human Neuroimaging, University College London, London, UK. <sup>6</sup>Department of Child and Adolescent Psychiatry, Psychotherapy and Psychosomatics, University of Würzburg, Würzburg, Germany. <sup>7</sup>Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland. <sup>8</sup>These authors contributed equally: Lorenz Deserno and Annette Horstmann. ✉email: lieneke.janssen@gmail.com

populations such as metamphetamine addiction, obsessive compulsive disorder, and binge eating disorder<sup>19,20</sup>, as well as in a general population sample reporting symptoms of the same disorders and of other eating disorders<sup>21</sup>. However, Voon et al.<sup>19</sup> did not find differences in model-based and model-free control between obese participants without binge eating disorder and non-obese control participants. The absence of an association between obesity and model-based or model-free control seems surprising, given the above-mentioned obesity-related performance differences in simple reinforcement learning tasks and outcome devaluation tasks, resembling more model-free and model-based control, respectively.

We propose two reasons why the study by Voon et al.<sup>19</sup> might have lacked power to detect obesity-related group differences in model-based and model-free control. First, rather subtle behavioural alterations are to be expected in obese individuals that are physically healthy. With a relatively low contrast in body mass index (BMI) between the obese and non-obese group (BMI [kg/m<sup>2</sup>]: obese:  $M = 31.49$ ,  $SD = 3.6$ ; non-obese:  $M = 23.54$ ,  $SD = 2.9$ ), and an average BMI for the obese group only slightly above the cutoff for obesity ( $> 30$  kg/m<sup>2</sup>), such behavioural alterations may be difficult to detect. Second, the relationship between BMI and model-based and model-free control may in fact be quadratic in nature, thus masking potential obesity-related differences. A quadratic relationship with degree of obesity has indeed been observed for reward sensitivity<sup>22</sup> and cognitive restraint of eating behaviour<sup>23</sup>. Furthermore, obesity may quadratically relate to alterations in striatal dopamine tone<sup>24</sup>. This is relevant because there is accumulating evidence that different measures and manipulations of dopamine transmission overall relate positively to model-based control as measured in the two-step task<sup>25–29</sup>.

In the current study, we aimed to address the two issues raised above by including (1) more highly obese individuals to boost the contrast between groups, and (2) an intermediate overweight group for more sensitivity to detect the existence of potential linear or quadratic relationships between weight status and behavioural control. The original two-step task was implemented to disentangle and directly compare the reliance on model-based and model-free control<sup>16,25,30</sup>. We hypothesized that weight status relates to the degree to which individuals rely on model-based and model-free learning, and that it may do so in a linear or quadratic manner.

## Materials and methods

**Participants.** The results reported in this study are based on data from 90 healthy right-handed participants in a wide BMI range (45 women; age [years]:  $M = 26.9$ ;  $SD = 3.6$ ; range 21–35; BMI [kg/m<sup>2</sup>]:  $M = 27.9$ ,  $SD = 6.4$ , range 18.4–47.6). Participants were recruited based on their BMI status, i.e., normal-weight [ $n(\text{women}) = 31(16)$ , BMI [kg/m<sup>2</sup>] = 18.5–24.9], overweight [ $n(\text{women}) = 29(14)$ , BMI [kg/m<sup>2</sup>] = 25–29.9] and obese [ $n(\text{women}) = 30(15)$ , BMI  $> 30$ ] (Table 1). Note that the reported data were acquired in two parts. Fifty-seven datasets were acquired as a part of several studies running in the department between October 2012 and August 2014. Data acquisition of overweight and obese participants was not completed at the time due to logistic reasons. To finally conclude the study, the remaining participants were tested between February and March 2018 ( $n = 37$ , for details see Supplemental Figure S1). Part of the reported data have previously been published in a study comparing relative reliance on model-based and model-free control to habit propensity in a slips-of-action task in specifically normal-weight women and men ( $n = 28$ )<sup>16</sup>. Participants were tested at the Department of Neurology of the Max Planck Institute for Human Cognitive and Brain Sciences (Leipzig, Germany) and received monetary compensation on an hourly basis, as well as a bonus based on their task performance (between 3 and 10€;  $M = 6.5€$ ,  $SD = 0.82$ ). All participants gave written consent prior to the study. The study was carried out in accordance with the Declaration of Helsinki and approved by the Ethics Committee at the University of Leipzig, Germany.

After having provided informed consent, weight and height of the participants was measured, followed by the two-step task (for details see “Experimental paradigm”). Participants were then asked to complete a number of self-report questionnaires—validated in German—for characterizing the sample: Beck’s Depression Inventory (BDI)<sup>31</sup> to assess possible depressive symptoms (cut-off for exclusion  $> 18$ , indicating possibility of moderate to severe depression), the Behavioural Inhibition System/Behavioural Activation System questionnaire (BIS/BAS)<sup>32,33</sup> to assess punishment and reward sensitivity, the Three-Factor Eating Questionnaire (TFEQ)<sup>34,35</sup> to assess eating behaviour in terms of cognitive restraint, disinhibition and hunger, the UPPS Impulsive Behaviour Scale<sup>36,37</sup> to assess impulsive behaviour in terms of Urgency, lack of Premeditation, lack of Perseverance, and Sensation seeking, and the Yale Food Addiction Scale (YFAS)<sup>38,39</sup> to assess symptoms that could be indicative of food addiction. Finally, participants performed several cognitive tests to examine their potential relation to performance on the task: the Viennese Matrices Test (VMT)<sup>40</sup> to assess non-verbal IQ. We also administered a computerized version of the Visual Paired Associates test of the Wechsler Memory Scale (VPA)<sup>41,42</sup> to assess visual short term memory. Participants were included if none of the following exclusion criteria applied: estimated non-verbal IQ ( $< 85$  based on the VMT), known metabolic disorders (e.g., diabetes), smoking, (history of) neurological, psychiatric, or eating disorders, symptoms of depression, drug or alcohol dependence, current pregnancy, and psychological treatment. In total 94 participants were tested of which three participants did not complete the experimental paradigm and one participant was excluded from analysis because of an estimated non-verbal IQ below 85.

**Experimental paradigm.** We administered a sequential decision making task<sup>16,25,30</sup>, in which participants were asked to make two subsequent decisions on each trial to earn a monetary reward (20 cents) or no reward (Fig. 1a). At the first stage, participants were asked to choose between two grey stimuli, which would bring them to one of two second-stage stimulus pairs (the green or yellow pair). One of the grey first-stage stimuli was connected commonly (70%) to the green and rarely (30%) to the yellow stimulus pair, and vice versa for the other grey stimulus (Fig. 1b). The first-stage stimuli and transition probabilities were fixed throughout the experiment. After selecting one of the two second-stage stimuli, participants either received the monetary reward or not (Fig. 1c). The probability of receiving reward for each of the four second-stage stimuli changed slowly and

|                                   | Normal-weight |            | Overweight   |            | Obese        |            | <i>p</i> | <i>Test-statistic</i> |
|-----------------------------------|---------------|------------|--------------|------------|--------------|------------|----------|-----------------------|
| n                                 | 31            |            | 29           |            | 30           |            |          |                       |
| sex (F:M)                         | 16:15         |            | 14:15        |            | 15:15        |            | ns       | 0.07 <sup>b</sup>     |
| Age                               | 26.9 (3.3)    | 21–34      | 26.0 (3.7)   | 21–35      | 27.8 (3.8)   | 22–34      | 0.166    | 1.8 <sup>c</sup>      |
| BMI (kg/m <sup>2</sup> )          | 21.6 (1.8)    | 18.4–24.8  | 26.9 (1.3)   | 25.1–29.9  | 35.4 (4.5)   | 30.2–47.6  | <0.001   | 79.1 <sup>d</sup>     |
| <b>Cognitive tests</b>            |               |            |              |            |              |            |          |                       |
| Non-verbal IQ <sup>a</sup>        | 119.9 (12.1)  | 95.0–136.5 | 117.3 (10.6) | 93.0–136.5 | 111.2 (16.4) | 85.0–136.5 | 0.084    | 5.0 <sup>d</sup>      |
| VPA score                         | 12 (3.9)      | 3–18       | 13.2 (3.1)   | 7–18       | 12.2 (3.2)   | 6–18       | 0.381    | 1.0 <sup>c</sup>      |
| <b>Self-report questionnaires</b> |               |            |              |            |              |            |          |                       |
| BDI                               | 3.6 (3.3)     | 0–14       | 4.9 (3.2)    | 0–11       | 6.8 (4.2)    | 0–17       | 0.003    | 11.5 <sup>d</sup>     |
| <b>BIS/BAS</b>                    |               |            |              |            |              |            |          |                       |
| BIS                               | 20.0 (3.4)    | 14–28      | 19.8 (4.3)   | 11–27      | 19.5 (4.0)   | 7–27       | 0.884    | 0.1 <sup>c</sup>      |
| BAS drive                         | 12.1 (2.1)    | 7–16       | 12.0 (1.8)   | 9–16       | 11.5 (1.7)   | 8–16       | 0.449    | 0.8 <sup>c</sup>      |
| BAS fun                           | 12.1 (1.8)    | 9–16       | 12.0 (1.8)   | 8–15       | 12.0 (1.9)   | 8–16       | 0.972    | 0.03 <sup>c</sup>     |
| BAS reward                        | 16.8 (2.0)    | 12–20      | 17.0 (2.1)   | 11–20      | 16.0 (2.0)   | 10–19      | 0.141    | 2.0 <sup>c</sup>      |
| <b>TFEQ</b>                       |               |            |              |            |              |            |          |                       |
| Restraint                         | 5.0 (3.2)     | 0–15       | 8.1 (4.6)    | 0–18       | 6.4 (4.7)    | 0–18       | 0.027    | 7.2 <sup>d</sup>      |
| Disinhibition                     | 4.9 (2.1)     | 0–9        | 6.3 (3.3)    | 2–15       | 8.3 (3.3)    | 3–16       | <0.001   | 16.9 <sup>d</sup>     |
| Hunger                            | 5.7 (3.3)     | 1–13       | 5.1 (4.1)    | 0–13       | 7.1 (3.4)    | 1–14       | 0.066    | 5.4 <sup>d</sup>      |
| <b>UPPS</b>                       |               |            |              |            |              |            |          |                       |
| Urgency                           | 26.8 (5.8)    | 15–42      | 25.4 (5.1)   | 17–36      | 27.7 (6.6)   | 13–39      | 0.314    | 1.2 <sup>c</sup>      |
| (lack of) Premeditation           | 22.2 (4.1)    | 12–31      | 22.7 (4.6)   | 16–36      | 22.3 (4.0)   | 12–29      | 0.904    | 0.1 <sup>c</sup>      |
| (lack of) Perseverance            | 20.1 (6.0)    | 12–44      | 19.4 (5.6)   | 10–34      | 21.3 (5.3)   | 12–34      | 0.438    | 0.8 <sup>c</sup>      |
| Sensation seeking                 | 31.8 (6.6)    | 18–44      | 31.6 (8.5)   | 17–48      | 28.0 (7.4)   | 14–40      | 0.090    | 2.5 <sup>c</sup>      |
| YFAS (#symptoms)                  | 0.8 (0.7)     | 0–2        | 1.3 (1.4)    | 0–7        | 1.9 (1.0)    | 0–4        | <0.001   | 17.3 <sup>d</sup>     |

**Table 1.** Group characteristics displaying mean (standard deviation) and range if not otherwise stated, followed by the test-statistic and *p*-value of group comparison for each measure. *N* number of participants; followed by the ratio of females to males; *VPA* visual paired associates test of the Wechsler Memory Scale; *BDI* Beck's Depression Inventory; *BIS/BAS* Behavioural Inhibition System/Behavioural Activation System; *TFEQ* Three-Factor Eating Questionnaire; *UPPS* Urgency, lack of Premeditation, lack of Perseverance, and Sensation seeking; *YFAS* Yale Food Addiction Scale. <sup>a</sup>Non-verbal IQ was calculated based on the Viennese Matrices Test (VMT). <sup>b</sup>Chi square test for frequency data (degrees of freedom: 2). <sup>c</sup>F-test with for normally distributed scores (degrees of freedom: 2.87). <sup>d</sup>Independent-Samples Kruskal–Wallis Test of distributions for non-normally distributed scores (degrees of freedom: 2).

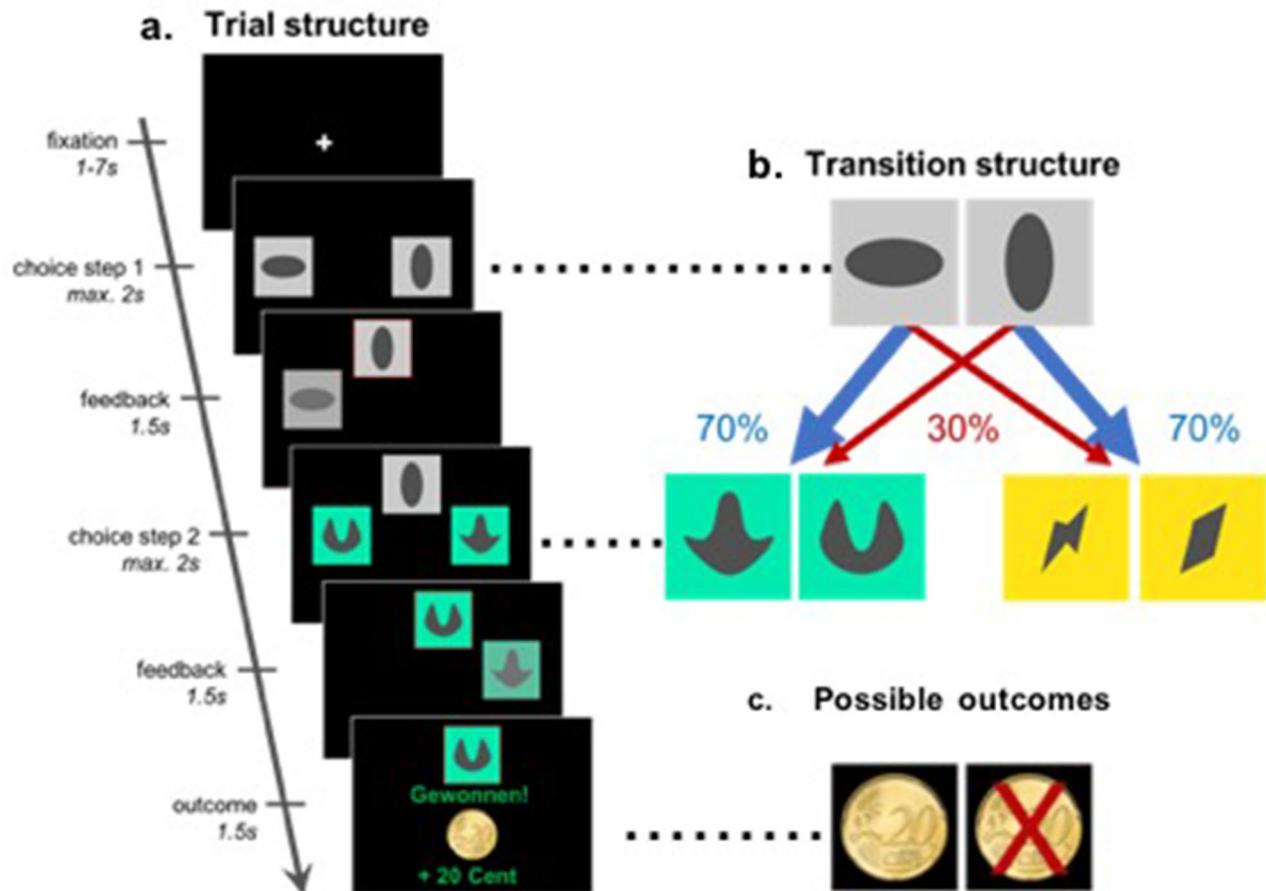
continuously according to Gaussian random walks to ensure continuous learning. The changes were kept consistent for all participants performing the experiment. Participants completed a total of 201 trials. Prior to the experiment, participants went through elaborate computer-based instructions and were then asked to explain the task including its first-stage transition probabilities to the experimenter. Open questions were addressed by the experimenter. The instructions included a detailed knowledge of common (70%) and rare (30%) transitions after first-stage choices, and the slowly changing probabilities after second-stage choices. After the instructions participants performed 56 training trials with a different set of stimuli. Participants were made aware that the height of their financial bonus depended on the accumulated reward in the task. The bonus was based on a randomly drawn subset of trials.

**Data analysis.** Calculation of first-stage stay probabilities on the two-step task, as well as computational modeling of participants' choice behaviour were performed using in-house scripts in Matlab (version 2017b, The MathWorks, Inc.). Statistical analyses of self-reported, behavioural, and computational data were run in R Studio (version 3.4.4., R Core Team, 2018<sup>43</sup>) and SPSS (version 24, IBM Corp., 2018). The R package ggplot2 was used to plot the results<sup>44</sup>.

Shapiro–Wilk's test of normality and Levene's test of equality of variance were ran for all group characteristics, including scores on self-reported questionnaires and neuropsychological tests, as well as for the accumulated reward (i.e., number of rewarded trials), raw stay probabilities (per condition), reaction times, and for the estimated model parameters.

The alpha level was set to 0.05 ( $\alpha = 0.05$ ) for all a priori analyses of interest. Note that for post hoc analyses, we did not correct for multiple comparisons as these results are exploratory and should be interpreted as such.

Partial  $\eta^2$  ( $\eta_p^2$ ) is reported as an effect size for all parametric univariate analyses because it meaningfully describes effects in a design in which multiple measures have been experimentally manipulated (as in the two-step task), and it yields very similar estimates as  $\eta^2$  for analyses that only include a between-group variable<sup>45,46</sup>. Note that  $\eta_p^2$  does not depend on the number of variables in the model and, thus, can be compared across studies.



**Figure 1.** The two-step task<sup>25,30</sup>. (a) Trial structure of an example trial with a rare transition, which allows for the dissociation of model-based and model-free control of behaviour. (b) Transition structure showing how each first-stage stimulus (grey) leads to one of the two second-stage stimulus pairs (green or yellow) in 70% of the trials (common, blue arrows) and to the other pair in 30% of the trials (rare, red arrows). (c) Possible outcomes (reward, no reward). Reward probability for the four second-stage stimuli varies throughout the task according to random walks to encourage continuous learning.

For non-parametric Kruskal–Wallis tests,  $\eta^2_H$  was calculated as follows:  $(H - k + 1)/(n - k)$ , with  $H$  reflecting the test statistic,  $k$  the number of groups, and  $n$  the total sample size<sup>47</sup>.

To check the robustness of our findings and rule out that any observed effect of group on behaviour could have been driven by age<sup>21,48,49</sup> or IQ<sup>16,21,50,51</sup> rather than weight status, we reran all models post hoc including age and non-verbal IQ as covariates of no interest.

**Characterization of the groups.** We tested for group differences in age and sex to confirm that the groups were well-matched. BMI was analysed to confirm the grouping of participants into normal-weight, overweight and obese participants. Group analysis of cognitive tests (including non-verbal IQ) and self-reported questionnaire data were run to further characterize the sample.

For normally distributed data (age, VPA score, BIS/BAS, UPPS), we ran a one-way ANOVA with between-subjects factor weight group for each measure. Upon violation of the assumption of normality or equality of variance (BMI, non-verbal IQ, BDI, TFEQ, YFAS symptom score), the Kruskal–Wallis test by ranks was performed. Sex distribution between groups was analysed using Chi-Square Test. Group differences were followed up by post hoc parametric (independent T-test) or nonparametric (Mann–Whitney U Test) pairwise comparisons.

**Raw behaviour according to first-stage stay probabilities.** Investigating the likelihood with which participants choose a first-stage stimulus depending on the previous trial type (Rewarded/Unrewarded, Common/Rare), gives an insight into how much they relied on model-based or model-free control. Therefore, we calculated first-stage stay probabilities as the proportion of trials in which participants chose the same first-stage stimulus as in the previous trial (coded as ‘stay’) for each of the conditions (Rewarded Common, Rewarded Rare, Unrewarded Common, Unrewarded Rare). We then analysed participants’ stay probabilities using ANOVA with the between-subject factor Group (Normal-weight, Overweight, Obese), and within-subject factors Reward (Rewarded, Unrewarded) and Transition (Common, Rare). Because the aim was to test for a three-way interaction and the group sizes are well balanced, type III sums of squares were calculated in this analysis.

A purely model-free agent relies on whether or not the previous trial was rewarded, irrespective of transition probability (Common/Rare). If rewarded, the previous first-stage choice should be repeated. If not, it may be better for the model-free agent to switch to the other first-stage stimulus. As a consequence, model-free control is reflected in a main effect of Reward. On the other hand, a purely model-based agent optimally relies both on reward and transition probability of the previous trial. A model-based agent will also stay with a previous first-stage choice when a common trial was rewarded, and switch when a common trial was not rewarded. However, the model-based agent differs in choice behaviour following rare trials. That is, in contrast to a purely model-free agent, a model-based agent can infer that when a rare trial was rewarded, reward probability on the current trial is higher if one chooses the other first-stage stimulus (switch), and vice versa for unrewarded rare trials (stay). Model-based control is therefore reflected in the interaction between Reward and Transition. Here, we were mainly interested in group differences in model-based and model-free control and thus focused on the Group  $\times$  Reward  $\times$  Transition interaction and Group  $\times$  Reward interaction on stay probabilities, respectively.

We hypothesized that the relationship between weight status and model-based or model-free control might be linear or quadratic in nature. To investigate the nature of these relationships, we next performed planned pairwise group comparisons on the Reward  $\times$  Transition interaction term [i.e., (Rewarded Common – Rewarded Rare) – (Unrewarded Common – Unrewarded Rare)] and on the main effect of Reward [i.e., (Rewarded Common + Rewarded Rare) – (Unrewarded Common + Unrewarded Rare)] on stay probabilities.

Finally, we ran two post hoc linear models (lm() from the R stats package): (1) on the Reward  $\times$  Transition interaction term, and (2) on the main effect of Reward to investigate the existence of a linear and quadratic relationship with BMI on a continuous scale. Both models included BMI and BMI<sup>2</sup> as orthogonal predictors.

**Computational modeling.** To investigate how participants' choices were affected by reward and transition probability throughout the experiment rather than in the previous trial alone, we computationally modeled choice behaviour. We implemented a hybrid of a model-free and model-based reinforcement algorithm as is described in detail in our previous work<sup>16,25</sup> and in the original paper<sup>30</sup>.

In short, the model-free algorithm (SARSA( $\lambda$ )) included a learning rate for each stage ( $\alpha_1, \alpha_2$ ) and a parameter  $\lambda$ , which allows the second stage prediction error to affect the next first-stage values (Q). The model-based algorithm learns values by planning forward and computes first-stage values by multiplying the value of the better second-stage option with the associated transition probabilities. Then, the model-free and model-based first-stage decision values are connected in the hybrid algorithm:

$$Q_{net}(s_A, a_j) = \omega Q_{MB}(s_A, a_j) + (1 - \omega) Q_{MF}(s_A, a_j)$$

where  $Q_{net}(s_A, a_j)$  denotes the decision value of the chosen stimulus  $a_j$  from the first stage stimulus pair  $s_A$ , and  $\omega$  captures the relative weighting of the model-based ( $Q_{MB}(s_A, a_j)$ ) and model-free algorithm ( $Q_{MF}(s_A, a_j)$ ). The weighting parameter  $\omega$  is the main parameter of interest and can take a value between 0 and 1. If  $\omega = 1$ , first-stage choices are purely controlled by model-based control, and if  $\omega = 0$ , they are purely controlled by model-free control. Note that at the second stage  $Q_{net} = Q_{MB} = Q_{MF}$ .

Finally, the decision values were transformed into action probabilities using the softmax function for  $Q_{net}$ :

$$P(a_{i,t} = a | s_{i,t}) = \frac{\exp(\beta_i [Q_{net}(s_{i,t}, a) + \rho \cdot rep(a)])}{\sum_{a'} \exp(\beta_i [Q_{net}(s_{i,t}, a') + \rho \cdot rep(a')])}$$

where  $\beta_i$  controls the stochasticity of choices at stage  $i = 1$  or 2, and repetition parameter  $\rho$  reflects choice perseveration at the first stage.

The model had a total of seven parameters that were bounded by transforming them to a logistic ( $\alpha_1, \alpha_2, \lambda, \omega$ ) or exponential ( $\beta_1, \beta_2$ ) distribution. To infer the maximum-a-posteriori estimate of each parameter for each subject, the (empirical) Gaussian prior distribution was set to the maximum-likelihood estimates given the data of all participants and then expectation-maximization was used<sup>52</sup>. We report the negative log-likelihood (–LL) as a measure of model fit.

We assessed group differences in  $\omega$  using ANOVA with between-group factor weight status. Planned pairwise comparisons were performed as part of the ANOVA or using Mann–Whitney U test as a nonparametric alternative. For each of these analyses, the alpha level was set at 0.05. Finally, we investigated the relationship between  $\omega$  and weight status on a continuous scale by running a post hoc linear regression model including BMI and BMI<sup>2</sup> as orthogonal predictors.

After having detected between-group differences on the model parameters' of interest, an important sanity check is whether the inferred parameters actually reproduce the observed behavioural data in terms of stay probabilities. To do so, we re-ran the model based on each individual's inferred parameters to generate data for each individual (1000 simulations per subject) and performed the original ANOVA.

We then ran simulation recovery analyses for the model to assess whether the model parameters captured the observed behavioural data. Based on the estimated parameters, we simulated choice behaviour on the task and investigated stay probabilities. The reported significant Group  $\times$  Reward  $\times$  Transition interaction was fully reproduced indicating that the model captured important aspects of the data (Supplemental Figure S2).

Finally, to confirm that the chosen hybrid model including  $\lambda$  was the best-fitting algorithm in this study, we compared the model to less complex models. To avoid inclusion of numerous combinations of parameters, we focus on models that capture distinct behaviour in this task by setting  $\omega$  to 1 or 0, and  $\lambda$  to 0 or fitting it as a free parameter. This gives four additional models: (1) a hybrid model without  $\lambda$  ( $\omega = 0$ ), (2) a model only including the model-based learning algorithm ( $\omega = 1$ ,  $\lambda$  can not be fitted), (3) a model only including the model-free learning

algorithm with  $\lambda$  ( $\omega=0$ ), and (4) the same model-free model without  $\lambda$  ( $\omega=0$ ). Integrated Bayesian Information Criterion (BIC) is reported for all models<sup>52</sup>.

## Results

**Characterization of the groups.** Table 1 summarizes the weight groups [normal-weight (NW), overweight (OW), and obese (OB)] in terms of age, sex and BMI, as well as in terms of their scores on the cognitive tests and self-report questionnaires. The groups were well matched on sex and age, and did not differ in visual short-term memory (VPA), or non-verbal IQ as measured on the Viennese Matrices Test (VMT). However, a trend-level group difference was observed for non-verbal IQ, with numerically higher IQ scores for the normal-weight and overweight relative to the obese group (Table 1). We did observe a group difference in the average number of depressive symptoms ( $KW(2)=11.5$ ,  $p=0.003$ ,  $\eta^2_H=0.11$ ) even though the scores are not clinically relevant in the current sample. This difference was driven by the obese participants having a higher symptom score relative to normal-weight, but not overweight, participants (post hoc pairwise comparisons: NW vs. OB,  $p=0.004$ ; OW vs. OB,  $p=0.137$ ; NW vs. OW,  $p=0.254$ ). Post hoc covariate analyses of behavioural and computational data controlling for BDI score did not change the primary effects of interest (see “Supplemental Materials” for statistics). The average number of food addiction symptoms also differed between the groups ( $KW(2)=17.3$ ,  $p<0.001$ ,  $\eta^2_H=0.18$ ), again, driven by a higher number of symptoms for obese relative to normal-weight, but not overweight, participants (post hoc pairwise comparisons: NW vs. OB,  $p<0.001$ ; OW vs. OB,  $p=0.159$ ; NW vs. OW,  $p=0.242$ ). In terms of self-reported eating behaviour (TFEQ) the groups differed in disinhibition ( $KW(2)=16.9$ ,  $p<0.001$ ,  $\eta^2_H=0.17$ ) and restraint ( $KW(2)=7.2$ ,  $p=0.027$ ,  $\eta^2_H=0.06$ ). Disinhibition scores were higher for obese relative to both normal-weight and overweight participants and somewhat higher for overweight relative to normal-weight participants (post hoc pairwise comparisons: NW vs. OB,  $p<0.001$ ; OW vs. OB,  $p=0.010$ ; NW vs. OW,  $p=0.076$ ). Restraint scores were highest for overweight participants and lower for normal-weight, but not obese participants (post hoc pairwise comparisons: NW vs. OB,  $p<0.375$ ; OW vs. OB,  $p=0.374$ ; NW vs. OW,  $p=0.013$ ). No other group differences were observed.

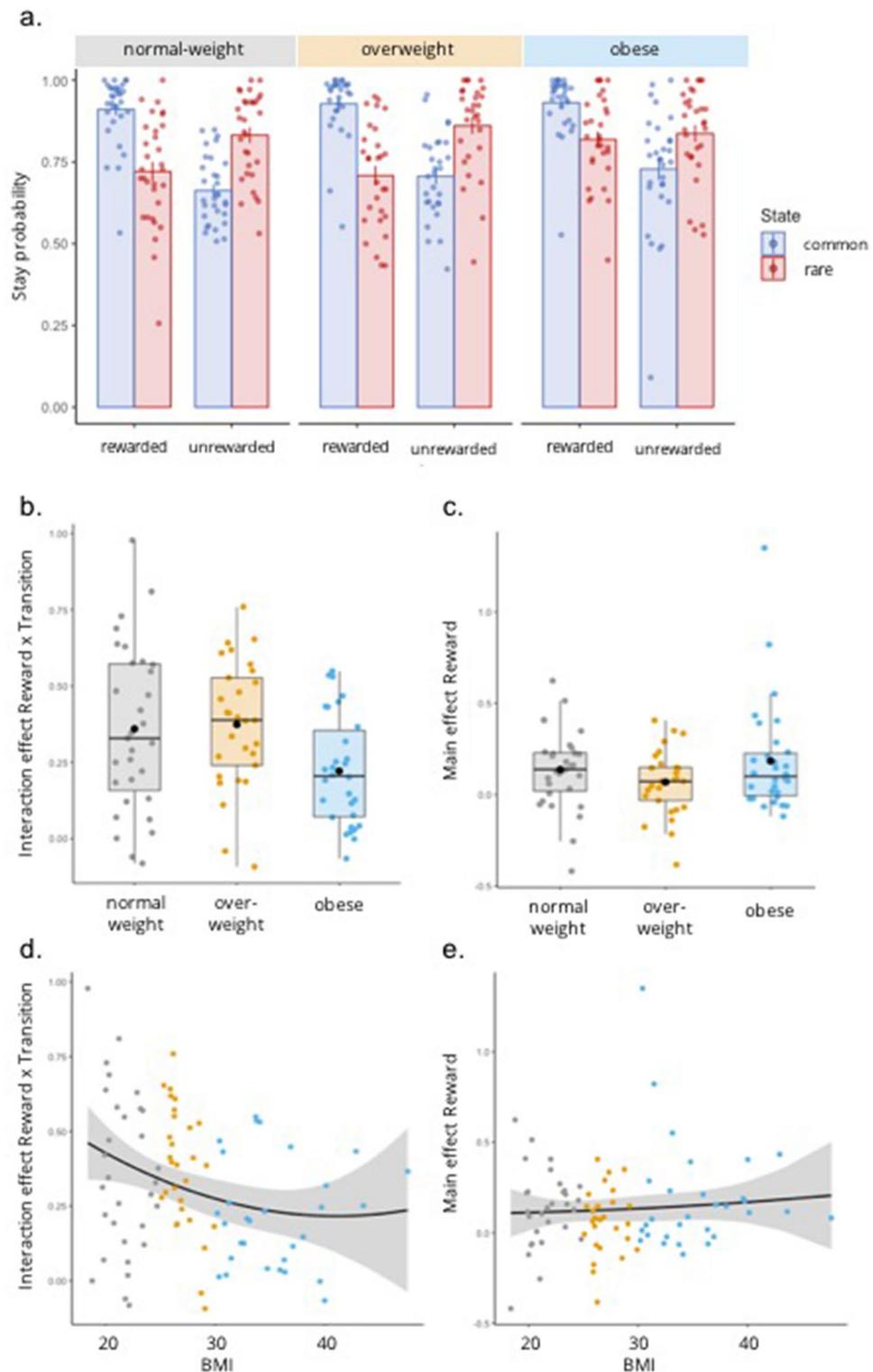
**Raw behaviour according to first-stage stay probabilities.** Analysis of stay probabilities (Fig. 2a) revealed that participants’ first-stage choices were significantly affected by reward (main effect Reward:  $F(1,87)=27.2$ ,  $p<0.001$ ,  $\eta_p^2=0.238$ ) as well as by the combination of reward and transition probability (interaction Reward  $\times$  Transition:  $F(1,87)=183.4$ ,  $p<0.001$ ,  $\eta_p^2=0.678$ ) on the previous trial. This is in line with previous research<sup>25,30</sup> and suggests that, across groups, the participants relied on both model-based and model-free choice strategies, respectively. Transition probability alone did not significantly affect participants’ first-stage choices (Transition:  $F(1,87)=3.4$ ,  $p=0.070$ ,  $\eta_p^2=0.037$ ).

The weight groups significantly differed in the use of a model-based choice strategy (Fig. 2b) as reflected by a significant three-way Group  $\times$  Reward  $\times$  Transition interaction on stay probabilities ( $F(2,87)=4.3$ ,  $p=0.017$ ,  $\eta_p^2=0.090$ ), but not in the use of a model-free choice strategy (Group  $\times$  Reward:  $F(2,87)=1.8$ ,  $p=0.174$ ,  $\eta_p^2=0.039$ , Fig. 2c). Planned comparisons of the Reward  $\times$  Transition interaction between groups showed that the three-way interaction was driven by a significantly higher interaction term for normal-weight relative to obese ( $p=0.017$ ) and for overweight relative to obese ( $p=0.010$ ) participants, whereas normal-weight and overweight participants did not differ from each other ( $p=0.817$ ).

We observed no Group  $\times$  Transition interaction ( $F(2,87)=1.2$ ,  $p=0.297$ ,  $\eta_p^2=0.028$ ), nor a main effect of Group ( $F(2,87)=1.7$ ,  $p=0.187$ ,  $\eta_p^2=0.038$ ) on stay probabilities. These results suggest that choices of obese participants relied relatively less on model-based control than those of normal-weight and overweight participants.

Post hoc simple effects analyses were performed to further investigate the three-way interaction on stay probabilities and revealed a striking difference between the groups. Interestingly, we observed a Group  $\times$  Reward interaction for rare ( $F(2,87)=4.2$ ,  $p=0.018$ ), but not common trials ( $F(2,87)<1$ ,  $p=0.497$ ). This in turn was driven by a simple main effect of Group on stay probabilities following rewarded rare trials ( $F(2,87)=4.6$ ,  $p=0.012$ ), but not unrewarded rare trials ( $F(2,87)<1$ ,  $p=0.688$ ). The simple effect of Group was also reflected in a Group  $\times$  Transition interaction for rewarded ( $F(2,87)=3.8$ ,  $p=0.026$ ), but not unrewarded trials ( $F(2,87)=2.4$ ,  $p=0.100$ ). Finally, pairwise group comparisons of rewarded rare trials showed that obese participants were more likely to stay with their previous first-stage choices when a rare trial had been rewarded relative to normal-weight ( $t(59)=-2.5$ ,  $p=0.014$ ) and overweight participants ( $t(57)=-2.9$ ,  $p=0.006$ ), with no difference between normal-weight and overweight participants ( $t(58)=0.3$ ,  $p=0.766$ ). This is of interest because it is participants’ behaviour following rare trials that allows us to dissociate model-based from model-free control. Increased staying after a rare rewarded trial hints at more model-free control, even though this effect was not sufficiently strong to come out as a significant interaction between Group and Reward. Nevertheless, it seems that the observed group difference in model-based control may in fact be driven by enhanced reliance on model-free computations (see “Discussion” for more).

Another means of probing the model-based control system in the framework of this task, is to investigate second-stage reaction times<sup>25,53</sup>. Since a model-based agent uses knowledge of the likelihood of the transition into a second-stage state, encountering an unexpected rare rather than an expected common transition should increase reaction times for choices at the second stage. In a post hoc analysis, we indeed observed a main effect of Transition on second-stage reaction times, which reflected significantly larger reaction times following a rare relative to a common transition ( $M(SD)_{\text{rare}}=992.3(137.5)$  ms  $>$   $M(SD)_{\text{common}}=731.1(101.2)$  ms,  $F(1,87)=366.2$ ,  $p<0.001$ ,  $\eta_p^2=0.808$ ). We did not, however, observe a significant Group by Transition interaction ( $F(2,87)<1$ ,  $p=0.411$ ,  $\eta_p^2=0.020$ ), which is in line with the above interpretation that the observed Group  $\times$  Reward  $\times$  Transition interaction on stay probabilities may not purely reflect a group difference in model-based control. The absence of the Group by Transition interaction could not be explained by general reaction times differences



**Figure 2.** Stay probabilities. **(a)** Average stay probabilities per condition for each group. Error bars represent  $\pm 1$  SEM. **(b)** On the group level, the use of a model-based choice strategy (i.e., the Reward  $\times$  Transition interaction term) was lower for obese relative to normal-weight and overweight participants, whereas **(c)** the use of a model-free choice strategy (i.e., the main effect of Reward) did not differ significantly between groups. The box plots in **(b)** and **(c)** show the median and interquartile range for each group, with the black dot denoting the mean. **(d)** On the continuous level, the Reward  $\times$  Transition interaction term was negatively related to BMI, with no additional significant quadratic relationship. **(e)** No linear or quadratic relationship was observed between BMI and the main effect of Reward. The scatter plots in **(d)** and **(e)** show the model fit (black line) and confidence interval (shaded) of the respective regression models with predictors BMI and BMI<sup>2</sup>. Individual data points are color-coded based on weight group for illustrative purpose.

|            |    |              | 25%   | 50%   | 75%   |                    |       |
|------------|----|--------------|-------|-------|-------|--------------------|-------|
| $\omega$   | NW | 0.66 (0.09)  | 0.58  | 0.68  | 0.72  | 5.3 <sup>a</sup>   | 0.007 |
|            | OW | 0.68 (0.09)  | 0.63  | 0.68  | 0.75  |                    |       |
|            | OB | 0.60 (0.11)  | 0.52  | 0.60  | 0.70  |                    |       |
| $\alpha_1$ | NW | 0.47 (0.17)  | 0.33  | 0.48  | 0.63  | < 1 <sup>b</sup>   | 0.867 |
|            | OW | 0.47 (0.21)  | 0.34  | 0.55  | 0.65  |                    |       |
|            | OB | 0.46 (0.29)  | 0.23  | 0.42  | 0.73  |                    |       |
| $\alpha_2$ | NW | 0.54 (0.23)  | 0.41  | 0.62  | 0.70  | < 1 <sup>a</sup>   | 0.560 |
|            | OW | 0.56 (0.19)  | 0.47  | 0.56  | 0.68  |                    |       |
|            | OB | 0.50 (0.21)  | 0.35  | 0.50  | 0.70  |                    |       |
| $\beta_1$  | NW | 7.7 (2.6)    | 5.3   | 7.8   | 9.0   | 1.6 <sup>b</sup>   | 0.447 |
|            | OW | 9.0 (4.0)    | 6.3   | 8.1   | 10.7  |                    |       |
|            | OB | 8.6 (3.2)    | 6.9   | 7.6   | 11.3  |                    |       |
| $\beta_2$  | NW | 4.3 (1.7)    | 2.8   | 4.1   | 5.4   | < 1 <sup>b</sup>   | 0.955 |
|            | OW | 4.0 (1.2)    | 3.2   | 4.2   | 4.8   |                    |       |
|            | OB | 5.2 (3.7)    | 3.3   | 4.0   | 5.4   |                    |       |
| $\lambda$  | NW | 0.53 (0.23)  | 0.36  | 0.54  | 0.71  | < 1 <sup>a</sup>   | 0.967 |
|            | OW | 0.53 (0.18)  | 0.38  | 0.56  | 0.69  |                    |       |
|            | OB | 0.55 (0.22)  | 0.41  | 0.55  | 0.69  |                    |       |
| $\rho$     | NW | 0.14 (0.05)  | 0.10  | 0.14  | 0.18  | 3.0 <sup>a,c</sup> | 0.057 |
|            | OW | 0.14 (0.04)  | 0.11  | 0.13  | 0.16  |                    |       |
|            | OB | 0.15 (0.06)  | 0.13  | 0.16  | 0.18  |                    |       |
| -LL        | NW | 175.7 (40.9) | 141.1 | 174.0 | 207.6 | 1.5 <sup>a</sup>   | 0.225 |
|            | OW | 169.8 (44.8) | 137.8 | 157.4 | 188.6 |                    |       |
|            | OB | 155.7 (50.7) | 128.3 | 157.9 | 193.3 |                    |       |

**Table 2.** Summary and group comparisons of all model parameters. *NW* normal-weight, *OW* overweight, *OB* obese. <sup>a</sup>F-test for normally distributed parameters (degrees of freedom: 2,87). <sup>b</sup>Independent-Samples Kruskal–Wallis Test of distributions for non-normally distributed parameters (degrees of freedom: 2). <sup>c</sup>One formal outlier was observed in the obese group and excluded from the analysis of this parameter (degrees of freedom: 2,86).

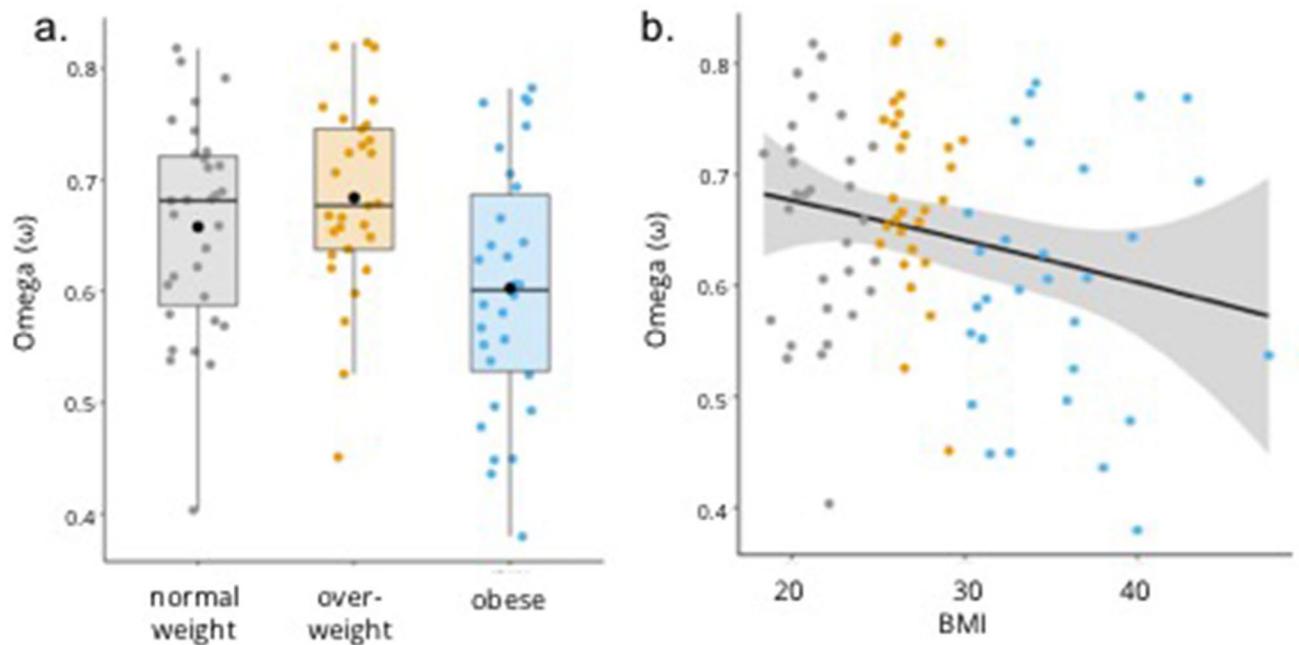
between the groups, as the groups did not differ in their RTs overall for either stage 1 ( $M(SD) = 656.7 (8.8) \text{ ms}$ ,  $F(2,87) < 1$ ,  $p = 0.847$ ,  $\eta_p^2 = 0.004$ ), or stage 2 decisions ( $M(SD) = 808.3 (9.6) \text{ ms}$ ,  $F(2,87) < 1$ ,  $p = 0.808$ ,  $\eta_p^2 = 0.005$ ).

Next, we addressed the question if reliance on model-based and model-free control related to obesity in a linear and/or quadratic manner. Because the traditional weight categories of normal-weight, overweight and obese individuals reflect unequal intervals in terms of BMI, we turned to BMI as a continuous variable, even though the study was designed for group-based analyses. We ran two linear regression models including BMI and BMI<sup>2</sup> as orthogonal predictors in each, and investigated their relationship with the (1) Reward  $\times$  Transition interaction term, and (2) the main effect of Reward on stay probabilities. BMI related negatively to the Reward  $\times$  Transition interaction term ( $\beta_{BMI} = -0.28$ ,  $p = 0.007$ ), but no additional quadratic relationship was observed ( $\beta_{BMI^2} = 0.10$ ,  $p = 0.319$ ) (Fig. 2d). Together, BMI and BMI<sup>2</sup> explained a significant proportion of variance in the effect of Reward and Transition on choice strategy (adjusted  $R^2 = 0.069$ ,  $F(2,87) = 4.3$ ,  $p = 0.017$ ). In line with the absence of a Group  $\times$  Reward effect on stay probabilities, we did not observe a linear or quadratic relationship between BMI and the main effect of Reward on stay probabilities ( $\beta_{BMI} = 0.08$ ,  $p = 0.463$ ;  $\beta_{BMI^2} = 0.01$ ,  $p = 0.892$ ) (Fig. 2e), nor did the model explain a significant proportion of variance (adjusted  $R^2 = -0.016$ ,  $F(2,87) = 0.3$ ,  $p = 0.756$ ). Note that a post hoc analysis results suggest that the linear relationship between BMI and the Reward  $\times$  Transition interaction term may be moderated by BDI score (see “Supplemental Materials” for statistics).

**Accumulated reward.** The accumulated reward (i.e., sum of rewarded trials) of participants was analyzed as a measure of overall performance. The groups did not differ in the sum of rewarded trials throughout the experiment ( $M = 97.2$ ,  $SD = 7.7$ ,  $F(2,87) = 1.6$ ,  $p = 0.209$ ,  $\eta_p^2 = 0.035$ ). The sum of rewarded trials also did not correlate to participants’ tendency to rely on model-based or model-free choice strategies in any of the measures of interest ( $p$ 's  $> 0.299$ ).

**Computational modeling of choice behaviour.** Computational modeling of behaviour allowed us to take into account participants’ choices throughout the experiment rather than only considering the effect of the previous trial. For a summary of all parameters and group comparisons, see Table 2.

The parameter  $\omega$  was of initial interest because it reflects participants’ relative reliance on model-based vs. model-free control. A purely model-based agent has an  $\omega$  of 1, whereas a purely model-free agent has an  $\omega$  of 0. As expected, we observed a significant group effect on  $\omega$  ( $F(2,87) = 5.3$ ,  $p = 0.007$ ,  $\eta_p^2 = 0.109$ ) (Fig. 3a). Planned



**Figure 3.** Relative reliance on model-based and model-free control (omega). **(a)** On the group level, omega was significantly lower for obese relative to normal-weight and overweight participants. The box plot reflects the median, interquartile range, and mean value (black dot) for each weight group. **(b)** On the continuous level, omega was negatively related to BMI, with no additional significant quadratic relationship. The scatter plot shows the model fit (black line) and confidence interval (shaded) of the regression model. Individual data points are color-coded based on weight group for illustrative purposes.

|            | Hybrid with $\lambda$ | Hybrid without $\lambda$ | MB       | MF with $\lambda$ | MF without $\lambda$ |
|------------|-----------------------|--------------------------|----------|-------------------|----------------------|
| All (n=90) | <b>32,140.72</b>      | + 273.55                 | + 402.40 | + 1089.91         | + 1733.09            |
| NW (n=31)  | <b>11,492.61</b>      | + 122.75                 | + 112.43 | + 423.32          | + 689.88             |
| OW (n=29)  | <b>10,696.63</b>      | + 62.30                  | + 85.73  | + 453.02          | + 630.55             |
| OB (n=30)  | <b>10,075.42</b>      | + 70.79                  | + 151.12 | + 195.87          | + 377.26             |

**Table 3.** Model comparison. Values reflect Bayesian Information Criterion (BIC) scores. Values in bold reflect the winning model when comparing the models across all participants, as well as in each group separately. Non-Bold scores reflect the difference with regard to the winning ‘full’ hybrid model with  $\lambda$ . NW normal-weight, OW overweight, OB obese, MB model-based, MF model-free.

comparisons showed that the group effect on  $\omega$  was driven by higher values for normal-weight relative to obese ( $t(59) = 2.1, p = 0.042$ ) and overweight relative to obese participants ( $t(57) = 3.1, p = 0.003$ ). Although overweight participants numerically had the highest  $\omega$  values, there was no statistical difference with normal-weight participants ( $t(58) = -1.1, p = 0.265$ ).

To investigate the nature of the relationship between  $\omega$  and weight on a continuous scale (i.e., BMI), we again ran a post hoc regression model including the linear term BMI and quadratic term BMI<sup>2</sup> as predictors. The linear term related negatively to values of  $\omega$  with lower values in individuals with a higher BMI ( $\beta_{BMI} = -0.23, p = 0.030$ ), whereas the quadratic term did not significantly add to the model ( $\beta_{BMI^2} = -0.005, p = 0.964$ ) (Fig. 3b). In total, the model explained 3.1% of variance in  $\omega$  (adjusted R<sup>2</sup> = 0.031,  $F(2,87) = 2.4, p = 0.093$ ), which reflects only a small effect of BMI on reliance on model-based vs. model-free control. Similar to post hoc analysis of the relationship between BMI and the interaction effect of Reward  $\times$  Transition on stay probabilities, the negative relationship between BMI and  $\omega$  may be moderated by BDI score (see ‘Supplemental Materials’ for statistics).

None of the other model parameters differed significantly between the groups (Table 2). This indicates that the groups did not differ in terms of first or second stage learning rates ( $\alpha_1, \alpha_2$ ), stochasticity of first or second stage choices ( $\beta_1, \beta_2$ ), the tendency to persevere independent of reward or transition ( $\rho$ ), the eligibility parameter ( $\lambda$ ), and importantly, how well the model fit participants’ data ( $-LL$ ).

Finally, to confirm that the chosen hybrid model including  $\lambda$  was the best-fitting algorithm in this study, we compared the model to four less complex models by setting  $\omega$  to 1 or 0, and  $\lambda$  to 0 or fitting it as a free parameter. Comparing the Bayesian Information Criterion (BIC) scores of the five models across the entire sample as well as in each group separately shows clear superiority for the ‘full’ hybrid model in each case (Table 3).

**Correcting for age and IQ.** To check the robustness of our findings and rule out that the observed group differences could be explained by age<sup>21,48,49</sup> or IQ<sup>16,21,50,51</sup> rather than weight status, we reran all models post hoc including age and non-verbal IQ as covariates of no interest. In case of nonparametric tests, the analyses were performed after having regressed out age and non-verbal IQ from the dependent variables using linear regression.

Adding the covariates did not change the results qualitatively—the outcomes were largely in line with the original analyses and suggest that weight status, over and above age and IQ, explains unique variance in the degree to which individuals rely on measures of model-based, and possibly model-free, control (see Supplemental Table S1 for a graphical overview of the outcomes of all analyses of interest). Notably, the reported group differences in model-based control, as observed in stay probabilities, and the relative reliance on model-based and model-free control, as reflected in the model parameter  $\omega$ , were relatively robust when correcting for age and non-verbal IQ. However, the pairwise comparison in model-based control between normal-weight and obese participants did not reach significance. Furthermore, on the continuous level we observed a similar negative relationship between BMI and model-based control (stay probabilities) (see “Supplemental Materials” for statistics).

## Discussion

The aim of this study was to investigate the relationship between weight status (i.e., normal-weight, overweight, and obese) and reliance on model-based and model-free control in the two-step task<sup>16,25,30</sup>. Our results indicate that obese participants relied less strongly on model-based control than overweight and—to a lesser extent—normal-weight participants, with no difference in performance between overweight and normal-weight participants. This was observed in group analysis of participants’ choice behaviour (i.e., stay probabilities), as well as in the continuous analysis where BMI negatively related to model-based choice behaviour. No quadratic relationship with BMI was observed. Furthermore, computational modeling of participants’ choices revealed a similar group difference in the weighting of model-based and model-free control (i.e.,  $\omega$ ) that was driven by less model-based control for obese relative to overweight and normal-weight participants. Secondary analyses, however, did not show group differences in the slowing of second-stage reaction times after rare transitions, as would be expected given the observed decrease in model-based choice behaviour in obesity.

Although seemingly contradictory, together these findings may in fact suggest that the observed obesity-related difference in model-based control is driven, in part, by enhanced reliance on model-free computations. This interpretation concurs with our post hoc simple effects analyses of stay probabilities, which revealed that the group difference in model-based control was driven by an increased inclination of obese (relative to normal-weight) to stay with their choice specifically after trials on which a rare transition led to reward. Rare trials are the trials of interest in this task, because performance following rare trials is used to dissociate model-based from model-free choices. Common trials, on the other hand, lead to the same decision in model-based and model-free agents. The group difference was only observed for rewarded, not unrewarded rare trials. We speculate that obese individuals may more easily fall back on model-free control, or in other words be more reactive after having been rewarded than normal-weight participants, whilst relying similarly on model-based control in the case of no reward. This speculative interpretation should be interpreted with care, as it has been shown that model-free control on this task can potentially be due to a misunderstanding of the model of the task<sup>54</sup>. Furthermore, the current task is not designed to address this subtle effect, which could explain why it was not reflected in a group difference in model-free control in the analysis of stay probabilities.

Our findings are in contrast to those of a previous study by Voon et al.<sup>19</sup> using the same paradigm. When comparing non-obese controls and obese participants with and without binge-eating disorder, Voon et al.<sup>19</sup> reported no difference in the weighting parameter  $\omega$  between obese participants *without* binge-eating disorder and non-obese controls, whereas  $\omega$  was on average lower for obese participants *with* binge-eating disorder relative to matched non-obese controls. Interestingly, our findings in healthy obese participants better match the previous findings in obese participants *with* binge-eating disorder. It should be noted however that  $\omega$ , and thus the reliance on model-based over model-free control, was much higher in the current study (mean (SD)  $\omega$ : 0.6 (0.11) vs. 0.3 (0.24), range 0–1). The discrepancy between the studies can be explained by several factors. First, the current study tested a more severely obese group than the Voon-study with a mean BMI of 35.4 kg/m<sup>2</sup> (SD 4.5) vs. 31.5 kg/m<sup>2</sup> (SD 3.6). In fact, in terms of BMI our sample was closer to the binge-eating group (mean BMI [kg/m<sup>2</sup>] 35.0, SD 5.6). It may thus be the case that the reported finding of a lower weighting parameter  $\omega$  in binge-eating disorder in the Voon-study can partially be explained by the severity of obesity. Alternatively, even though no psychiatric conditions were reported, the obese participants in our sample might unbeknownst fulfill criteria for binge-eating disorder or have other co-morbidities, as we did not conduct a full psychiatric screening. We did, however, observe a group difference in self-reported depressive symptom score as assessed by Beck’s Depression Inventory (BDI), with higher—but subclinical—scores for the obese relative to the normal-weight and overweight group. Post hoc analyses showed that variation in BDI scores could not explain the observed group difference in reliance on model-based control. On the continuous level, BDI score did seem to moderate the negative relationship between model-based control and BMI, with a stronger negative relationship for higher BDI scores. It should be noted that the results of the post hoc analyses have to be interpreted with caution because BDI score was not systematically sampled and the use of the BDI as a continuous measure of depressive symptoms in obesity is criticized. The BDI includes both non-somatic and somatic items. High scores on the somatic items (e.g., fatigue, sleep disturbance, body image) may either reflect true depressive symptoms or they are instead related to individuals’ obesity. Second, we included an intermediate weight group for increased sensitivity to detect group differences and potential quadratic effects that might otherwise remain uncovered. The group difference in model-based control in the current study was indeed mostly driven by the difference between overweight and obese participants. We therefore recommend that cognitive studies of obesity

should include a wide BMI range, preferably also sampling severe to morbid obesity to assess for quadratic relationships, and to carefully disentangle between contributions of weight status and compulsive measures such as binge-eating symptoms.

The observed difference in reliance on model-based control in obesity generally concurs with previous outcome devaluation studies in relation to obesity that found reduced goal-directed control<sup>13,14</sup>. Goal-directed and model-based control are often equated<sup>11</sup> and have been found to relate, albeit weakly<sup>15–17</sup>. However, the concepts measured in the two types of tasks do not reflect the exact same constructs. Whereas the two-step task is designed to dissociate model-based and model-free control, it is difficult to disentangle reliance on goal-directed and habitual control in outcome devaluation paradigms in humans. In fact, for the current version of the two-step task—with reliance on model-based vs. model-free choice strategy *not* affecting overall outcome—one could paradoxically speculate that those who rely more strongly on model-free control are putatively even more efficient. For model-based control to be a more sensible strategy than model-free control, it should pay off to spend the extra cognitive resources associated with it<sup>55</sup>. That participants indeed follow this strategy was recently confirmed in a similar sequential decision-making task in which the incentive size was manipulated: model-based control indeed increased with larger incentives in a heterogeneous nonpatient population<sup>56</sup>. Furthermore, goal-directed and habitual control may be organized hierarchically rather than in parallel. That is, the goal-directed system may benefit from habits in goal-pursuit and thus rely on the habit system<sup>57</sup>, and the habit system may affect what goals are selected and pursued by the goal-directed system<sup>58</sup>. Empirical evidence for the existence of such hierarchies comes from a new generation of sequential decision-making tasks<sup>59–61</sup>. It will be relevant for future studies to focus on habitual goal-selection in the context of obesity, as has been suggested for addiction and other disorders of compulsivity<sup>58</sup>, and investigate if it relates more closely to maladaptive eating behaviour in daily life.

The current study has several limitations. First, the dataset was collected in two parts with a sampling bias in terms of group and sex (see Supplemental Figure S1). Due to this bias we could not meaningfully account for sex and sample (2012–2014 vs. 2018) as covariates of no interest, because variance explained by sample and weight group or sample and sex cannot be disentangled in our design<sup>62</sup>. However, the task was identical in both sampling periods and administered in very similar lab spaces within the department. More importantly, extensive computerized instructions were implemented to minimize variability in performance due to differences in instructions between experimenters. We are therefore fairly confident that the observed group differences are not confounded by sampling period. Second, as emphasized above, the observed group differences are subtle with modest effect sizes and await replication. We speculate that these differences may be more pronounced when taking into account participants' diet rather than obesity. Rodent studies suggest that rather than obesity, the intake of high fat and/or sugar diets may better predict alterations in dopamine-transmission<sup>63–68</sup>. We expect these changes to be at the heart of the maladaptive behavioural control in obesity<sup>24</sup> and there is accumulating evidence that different measures and manipulations of dopamine transmission overall related positively to model-based control as measured in the two-step task<sup>25–29</sup>. Whether diet rather than obesity relates to maladaptive behavioural control needs to be addressed in further studies. A third limitation is that, although the continuous analyses converge with the observed group differences in model-based control and strengthens the conclusion that obesity is indeed associated with altered reliance on model-based vs. model-free control, the design of the current study was not optimal for this type of analysis. BMI was not equidistributed across the complete sample due to the group-based recruitment-strategy. Hence, the current study might have been underpowered to robustly show true effects between BMI and behavioural control strategies on a continuous level. Another reason for interpreting the reported relationship between BMI and behavioural control on a continuous level with care is the low retest reliability of the task, as has recently been shown in a large-scale investigation of self-regulation paradigms<sup>69</sup>. For the investigation of individual differences in task performance within groups, other variables such as reaction time and latent variables from drift diffusion modeling could give a more reliable estimate of behavioural control<sup>70</sup>. Following Hedge et al.<sup>71</sup>, the poor retest reliability that is the result of low between-subject variability does not negate the observed group differences. In fact, low-between subject variability is required to observe reliable group differences in task performance. Despite these limitations, the findings from our two independent analysis approaches did converge. That is, analysis of raw choice behaviour in terms of stay probabilities and of the model parameter  $\omega$  both point to alterations in the reliance on model-based vs. model-free control in obesity. Simulation recovery analysis of the parameter estimates of the computational models further strengthened our confidence in the observed findings, because it recovered the observed three-way interaction between group, reward and transition probability on stay probabilities.

In conclusion, we found evidence for a relationship between the degree of obesity and reliance on model-based and model-free control relative to overweight and normal-weight participants, which may in fact be linear rather than quadratic in nature. Obesity, on the group-level, was associated with relatively lower model-based control compared to normal-weight and overweight, which was driven by an increased inclination of obese (relative to normal-weight) to stay with their choice specifically after trials on which a rare transition led to reward. Together, our findings suggest that it may be the combination of decreased model-based and increased model-free control in this task that characterizes the obese group. Whether or not the observed effects are dopamine-mediated, as hypothesized, remains an open question that warrants further investigation, for example, by pharmacologically manipulating dopamine transmission, or investigating the interaction between BMI and individual differences in dopamine transmission in terms of genetic or epigenetic variation.

### Data availability

The datasets analysed during the current study are available from the corresponding author on reasonable request.

Received: 19 July 2019; Accepted: 8 December 2020

Published online: 31 December 2020

## References

- Horstmann, A. It wasn't me; it was my brain—Obesity-associated characteristics of brain circuits governing decision-making. *Physiol. Behav.* **176**, 125–133 (2017).
- Volkow, N. D., Wise, R. A. & Baler, R. The dopamine motive system: Implications for drug and food addiction. *Nat. Rev. Neurosci.* **18**, 741 (2017).
- Lowe, C. J., Reichelt, A. C. & Hall, P. A. The prefrontal cortex and obesity: A health neuroscience perspective. *Trends Cogn. Sci.* **23**, 349–361 (2019).
- García-García, I. *et al.* Reward processing in obesity, substance addiction and non-substance addiction. *Obes. Rev.* **15**, 853–869 (2014).
- Kroemer, N. B. & Small, D. M. Fuel not fun: Reinterpreting attenuated brain responses to reward in obesity. *Physiol. Behav.* **162**, 37–45 (2016).
- Coppin, G., Nolan-Poupart, S., Jones-Gotman, M. & Small, D. M. Working memory and reward association learning impairments in obesity. *Neuropsychologia* **65**, 146–155 (2014).
- van den Akker, K., Schyns, G. & Jansen, A. Altered appetitive conditioning in overweight and obese women. *Behav. Res. Ther.* **99**, 78–88 (2017).
- Meemken, M. T., Kube, J., Wickner, C. & Horstmann, A. Keeping track of promised rewards: Obesity predicts enhanced flexibility when learning from observation. *Appetite* **131**, 117–124 (2018).
- Mathar, D., Neumann, J., Villringer, A. & Horstmann, A. Failing to learn from negative prediction errors: Obesity is associated with alterations in a fundamental neural learning mechanism. *Cortex* **95**, 222–237 (2017).
- Kube, J. *et al.* Altered monetary loss processing and reinforcement-based learning in individuals with obesity. *Brain Imaging Behav.* **12**, 1431–1449 (2018).
- Dolan, R. J. & Dayan, P. Goals and habits in the brain. *Neuron* **80**, 312–325 (2013).
- Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704 (2005).
- Horstmann, A. *et al.* Slave to habit? Obesity is associated with decreased behavioural sensitivity to reward devaluation. *Appetite* **87**, 175–183 (2015).
- Janssen, L. K. *et al.* Loss of lateral prefrontal cortex control in food-directed attention and goal-directed food choice in obesity. *Neuroimage* **146**, 148–156 <https://doi.org/10.1016/j.neuroimage.2016.11.015> (2017).
- Gillan, C. M., Otto, A. R., Phelps, E. A. & Daw, N. D. Model-based learning protects against forming habits. *Cogn. Affect. Behav. Neurosci.* **15**, 523–536 (2015).
- Sjoerds, Z. *et al.* Slips of action and sequential decisions: A cross-validation study of tasks assessing habitual and goal-directed action control. *Front. Behav. Neurosci.* **10**, 234 (2016).
- Friedel, E. *et al.* Devaluation and sequential decisions: Linking goal-directed and model-based behavior. *Front. Hum. Neurosci.* **8**, 587 (2014).
- Byrne, K. A., Otto, A. R., Pang, B., Patrick, C. J. & Worthy, D. A. Substance use is associated with reduced devaluation sensitivity. *Cogn. Affect. Behav. Neurosci.* **19**, 40–55 (2019).
- Voon, V. *et al.* Disorders of compulsivity: A common bias towards learning habits. *Mol. Psychiatry* **20**, 345–352 (2015).
- Voon, V. *et al.* Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. *Transl. Psychiatry* **5**, e670 (2015).
- Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A. & Daw, N. D. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife* **5**, e11305 (2016).
- Davis, C., Strachan, S. & Berkson, M. Sensitivity to reward: implications for overeating and overweight. *Appetite* **42**, 131–138 (2004).
- Dietrich, A., Federbusch, M., Grellmann, C., Villringer, A. & Horstmann, A. Body weight status, eating behavior, sensitivity to reward/punishment, and gender: Relationships and interdependencies. *Front. Psychol.* **5**, 1073 (2014).
- Horstmann, A., Fenske, W. K. & Hankir, M. K. Argument for a non-linear relationship between severity of human obesity and dopaminergic tone. *Obes. Rev.* **16**, 821–830 (2015).
- Deserno, L. *et al.* Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proc. Natl. Acad. Sci.* **112**, 1595 LP – 1600 (2015).
- Sharp, M. E., Foerde, K., Daw, N. D. & Shohamy, D. Dopamine selectively remediates 'model-based' reward learning: A computational approach. *Brain* **139**, 355–364 (2016).
- Doll, B. B., Bath, K. G., Daw, N. D. & Frank, M. J. Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *J. Neurosci.* **36**, 1211–1222 (2016).
- Wunderlich, K., Smittenaar, P. & Dolan, R. J. Dopamine enhances model-based over model-free choice behavior. *Neuron* **75**, 418–424 (2012).
- Kroemer, N. B. *et al.* L-DOPA reduces model-free control of behavior by attenuating the transfer of value to action. *Neuroimage* **186**, 113–125 (2019).
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
- Kühner, C., Bürger, C., Keller, F. & Hautzinger, M. Reliabilität und Validität des revidierten Beck-Depressionsinventars (BDI-II). *Nervenarzt* **78**, 651–656 (2007).
- Carver, C. S. & White, T. L. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *J. Pers. Soc. Psychol.* **67**, 319–333 (1994).
- Strobel, A., Beauducel, A., Debener, S. & Brocke, B. Eine deutschsprachige version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differ. und Diagnostische Psychol.* **22**, 216–227 (2001).
- Stunkard, A. J. & Messick, S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J. Psychosom. Res.* **29**, 71–83 (1985).
- Pudel, V. & Westenhöfer, J. *Fragebogen zum essverhalten (FEV): handanweisung* (Verlag für Psychologie Hogrefe, 1989).
- Whiteside, S. P. & Lynam, D. R. The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Pers. Individ. Dif.* **30**, 669–689 (2001).
- Schmidt, R. E., Gay, P., d'Acremont, M. & Van der Linden, M. A German adaptation of the UPPS impulsive behavior scale: Psychometric properties and factor structure. *Swiss J. Psychol.* **67**, 107–112 (2008).
- Gearhardt, A. N., Corbin, W. R. & Brownell, K. D. Preliminary validation of the Yale food addiction scale. *Appetite* **52**, 430–436 (2009).
- Meule, A., Vögele, C. & Kübler, A. Deutsche Übersetzung und Validierung der Yale food addiction scale. *Diagnostica* **58**, 115–126 (2012).

40. Formann, A. K. & Piswanger, K. *Wiener Matrizen-Test: WMT; ein Rasch-skaliertes sprachfreies Intelligenztest; Manual; mit englisch- und französischsprachigen Instruktionen im Anhang* (Beltz Test Ges., Chicago, 1979).
41. Wechsler, D. *Wechsler Memory Scale-revised. Manual* (Psychological Corporation, Chicago, 1987).
42. von Aster, M. N. A. & Horn, R. Wechsler Intelligenztest für Erwachsene WIE. Übersetzung und Adaption der WAIS-III von David Wechsler [German adaption of the WAIS-III] (2006).
43. Team, R. C. R: A language and environment for statistical computing. (2013).
44. Wickham, H. *ggplot2: elegant graphics for data analysis* (Springer, New York, 2016).
45. Baguley, T. Standardized or simple effect size: What should be reported?. *Br. J. Psychol.* **100**, 603–617 (2011).
46. Levine, T. R. & Hullett, C. R. Eta squared, partial eta squared, and misreporting of effect size in communication research. *Hum. Commun. Res.* **28**, 612–625 (2002).
47. Cohen, B. H. *Explaining Psychological Statistics* (Wiley, New York, 2008).
48. Eppinger, B., Walter, M., Heekeren, H. R. & Li, S.-C. Of goals and habits: Age-related and individual differences in goal-directed decision-making. *Front. Neurosci.* **7**, 253 (2013).
49. Eppinger, B., Schuck, N. W., Nystrom, L. E. & Cohen, J. D. Reduced striatal responses to reward prediction errors in older compared with younger adults. *J. Neurosci.* **33**, 9905–9912 (2013).
50. Otto, A. R., Gershman, S. J., Markman, A. B. & Daw, N. D. The curse of planning: Dissecting Multiple reinforcement-learning systems by taxing the central executive. *Psychol. Sci.* **24**, 751–761 (2013).
51. Otto, A. R., Skatova, A., Madlon-Kay, S. & Daw, N. D. Cognitive control predicts use of model-based reinforcement learning. *J. Cogn. Neurosci.* **27**, 319–333 (2014).
52. Huys, Q. J. *et al.* Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput. Biol.* **7**(4), e1002028 (2011).
53. Decker, J. H., Otto, A. R., Daw, N. D. & Hartley, C. A. From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychol. Sci.* **27**, 848–858 (2016).
54. da Silva, C. F. & Hare, T. A. Humans are primarily model-based and not model-free learners in the two-stage task. *bioRxiv* <https://doi.org/10.1101/682922> (2019).
55. Kool, W., Cushman, F. A. & Gershman, S. J. When does model-based control pay off?. *PLOS Comput. Biol.* **12**, e1005090 (2016).
56. Patzelt, E. H., Kool, W., Millner, A. J. & Gershman, S. J. Incentives boost model-based control across a range of severity on several psychiatric constructs. *Biol. Psychiatry* **85**(5), 425–433 (2019).
57. Wood, W. & Neal, D. T. A new look at habits and the habit-goal interface. *Psychol. Rev.* **114**, 843–863 (2007).
58. Daw, N. D. Of goals and habits. *Proc. Natl. Acad. Sci.* **112**, 13749–13750 (2015).
59. Dezfouli, A. & Balleine, B. W. Actions, action sequences and habits: Evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Comput. Biol.* **9**, e1003364 (2013).
60. Keramati, M., Smittenaar, P., Dolan, R. J. & Dayan, P. Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. *Proc. Natl. Acad. Sci.* <https://doi.org/10.1073/pnas.1609094113> (2016).
61. Cushman, F. & Morris, A. Habitual control of goal selection in humans. *Proc. Natl. Acad. Sci.* **112**, 13817 LP – 13822 (2015).
62. Miller, G. A. & Chapman, J. P. Misunderstanding analysis of covariance. *J. Abnorm. Psychol.* **110**, 40–48 (2001).
63. Friend, D. M. *et al.* Basal ganglia dysfunction contributes to physical inactivity in obesity. *Cell Metab.* **25**, 312–321 (2017).
64. Alsiö, J. *et al.* Dopamine D1 receptor gene expression decreases in the nucleus accumbens upon long-term exposure to palatable food and differs depending on diet-induced obesity phenotype in rats. *Neuroscience* **171**, 779–787 (2010).
65. Cone, J. J., Chartoff, E. H., Potter, D. N., Ebner, S. R. & Roitman, M. F. Prolonged high fat diet reduces dopamine reuptake without altering DAT gene expression. *PLoS ONE* **8**, e58251 (2013).
66. Baladi, M. G., Horton, R. E., Owens, W. A., Daws, L. C. & France, C. P. Eating high fat chow decreases dopamine clearance in adolescent and adult male rats but selectively enhances the locomotor stimulating effects of cocaine in adolescents. *Int. J. Neuropsychopharmacol.* **18**(7) (2015).
67. Li, Y. *et al.* High-fat diet decreases tyrosine hydroxylase mRNA expression irrespective of obesity susceptibility in mice. *Brain Res.* **1268**, 181–189 (2009).
68. Davis, J. F. *et al.* Exposure to elevated levels of dietary fat attenuates psychostimulant reward and mesolimbic dopamine turnover in the rat. *Behav. Neurosci.* **122**, 1257–1263 (2008).
69. Enkavi, A. Z. *et al.* Large-scale analysis of test-retest reliabilities of self-regulation measures. *Proc. Natl. Acad. Sci.* **116**, 5472 LP – 5477 (2019).
70. Shahar, N. *et al.* Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Comput. Biol.* **15**, e1006803 (2019).
71. Hedge, C., Powell, G. & Sumner, P. The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behav. Res. Methods* **50**, 1166–1186 (2018).

## Acknowledgements

The authors thank Anja Dietrich, Tilmann Wilbertz, and Sarah Kusch for their help with data acquisition, and Zsuzsika Sjoerds for the helpful discussion about the data. The authors are also grateful to Nils Kroemer, Ying Lee, Kathleen Wiencke, and the O'BRAIN Lab for their constructive feedback regarding data-analysis. This work was supported by the Federal Ministry for Education and Research, Germany, FKZ: 01EO1501 (LKJ, LD, AH) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (Project number 209933839-SFB 1052), SFB 1052 “Obesity Mechanisms”, subproject A05 (AH).

## Author contributions

A.H., L.D., and F.S., designed the experiment. F.M. and L.J. set up the study. F.M. collected data. L.J., L.D., and F.M. analysed the data. L.J., L.D., and A.H. wrote the manuscript. All authors read and approved the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-020-79929-0>.

**Correspondence** and requests for materials should be addressed to L.K.J.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020