

## SUPPLEMENTARY MATERIALS

### **Study1: Methods**

#### *Participants*

Participants were, on average, 53.42 years of age ( $SD = 10.57$ , range = 21 – 68 years). The racial/ethnic makeup of the sample was predominately White/Caucasian ( $n = 59$ , 92.19%), with five participants identifying as racial and ethnic minorities (Black = 1, 1.56%; Asian = 1, 1.56%; Hispanic = 1, 1.56%; Native Hawaiian or Pacific Islander = 1, 1.56%; Multiracial = 1, 1.56%). On average, participants were overweight, as defined by having a body mass index (BMI) above 25 (Williams, Mesidor, Winters, Dubbert, & Wyatt, 2015): participants' average weight at baseline was 190.18 pounds ( $M=190.18$ ,  $SD = 41.27$ ), with an average body mass index (BMI) of 33 ( $SD = 7.16$ ). Most participants had previous meditation experience ( $n = 38$ , 59.37%), and 84.2% among those ( $n = 32$ ) reported practice of daily meditation. The majority of participants identified sweet food as the category of foods that they crave ( $n = 45$ , 70.31%), followed by salty types of foods ( $n = 19$ , 29.69%). Participant characteristics are displayed in Table 1.

#### *Recruitment Procedure*

Interested individuals were considered eligible if they (a) were 18 to 70 years of age; (b) assigned female at birth and identify as a cisgender female; (c) able to read and speak English; (d) experienced food (salty or sweet) cravings as well as overate these foods (i.e. responded 'yes' when asked 'do you find yourself eating more than you'd like of a particular food or category of foods?') at least 4 times per week; (e) owned a smartphone; (e) resided within

driving distance of a city in northeast United States where the study assessment is conducted, and (f) indicated desire for eating behavior change. Exclusion criteria included (a) a current diagnosis of eating disorder; (b) a current strict diet (e.g., paleo, keto, vegan, calorie restriction); (c) active use of insulin; (d) current pregnancy; and (e) previous use of the mHealth application.

### *Mindful Eating Craving Tool*

To use the mindful eating craving tool, participants were instructed to click “craving tool” on the app’s dashboard whenever they experienced a food craving. This would initiate the start of the guided mindful eating exercise described as follows (an example of screenshots taken at each stage of a prototypical craving tool use is illustrated in Figure 2A). As a first step (Figure 2 A- Screenshot1), participants were invited to: identify and think of the food they are craving, simulate an experience of eating this food, and pay mindful attention to the phenomenological qualities experienced by this eating simulation (bodily sensations, emotions, and thoughts). After having completed this mindful eating simulation exercise, participants were then asked to rate the intensity of their craving (relative to what their baseline level of craving was before having completed the mindful eating simulation), and asked whether they wished to eat the craved food or not (Figure 2. A-Screenshot2). If participants chose not to eat, the exercise resumed, but if they opted to eat, they were prompted to rate the quantity of food eaten (Figure 2.A-Screenshot3). Next, participants were invited to pay mindful attention to bodily sensations, emotions, and thoughts experienced after having eaten, and asked to rate their level of contentment experienced from having eaten (Figure 2. A-Screenshot4). Following this, the participant was given the option to either resume the mindful eating exercise, or to receive another check-in five minutes later. If the individual opted to check-in 5 minutes later, they were directed to focus on present-moment

bodily sensations/emotions/thoughts and rate their contentment levels experienced from having eaten. The participant was then given the final option to resume the use of the mindful eating craving tool, or to go through the same check-in procedure 15 minutes later. A flow diagram (Figure 2B) conceptually displays the sequence of events, psychological processes or behaviors involved at each stage of a prototypical craving tool use.

### *Questionnaires Assessing Self-reported Maladaptive Eating Behaviors*

Self-reported maladaptive eating behaviors were assessed using the three following questionnaires. First, the Salzburg Stress Eating Scale (SSES) is a 10-item self-report measure of the extent to which an individual eats in response to stress (Meule, Reichenberger, & Blechert, 2018). Participants were instructed to respond to the items describing situations of perceived stress (e.g., “When I feel nervous and stressed,” “When I feel difficulties have been piling up so high that I cannot overcome them”) and their corresponding eating behaviors on a Likert Scale from 1 (*I eat much less than usual*) to 5 (*I eat much more than usual*). Higher scores indicate eating more than usual in response to stress and lower scores indicate eating less than usual in response to stress. Cronbach’s  $\alpha$  in our sample at baseline and post-intervention were .88 and .93, respectively.

Second, the Reward Based Eating Drive (RED) Scale is a self-report measure of reward-based eating (Epel et al., 2014). This questionnaire assesses constructs of 1) lack of control over cravings, 2) lack of satiety after eating and 3) over-preoccupation with food. Participants were asked to rate on 13 items regarding their eating behaviors related to these dimensions on a Likert scale from 0 (*Strongly disagree*) to 4 (*Strongly agree*). Higher scores indicate higher levels of reward-based eating (e.g., “When I start eating, I just can’t seem to stop,” “Food is always on my mind”). Cronbach’s alphas in our sample at baseline and post-intervention were .88 and .91, respectively.

Last, the Food Craving Questionnaire- Trait, Reduced (FCQ-T-r) consists of 15-items measuring the food craving experience patterns as a general tendency or as a psychological trait/characteristic (Meule, Hermann, & Kübler, 2014). This scale focuses on essential cognitive and behavioral aspects of food craving experiences such as thinking about food, intending to eat food, and losing control over food intake. Participants rated items on a 5-point Likert scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*). Sample items include “If I eat what I am craving, I often lose control and eat too much,” “I find myself preoccupied with food,” and “I crave foods when I feel bored, angry, or sad.” Cronbach’s  $\alpha$  in our sample at S1 and S2 were .91 and .96, respectively.

#### *Data Analysis: Computational Model Description*

A Rescorla-Wagner model of reinforcement learning (Rescorla and Wagner, 1972) was used to estimate participants’ expected reward values assigned to their eating behaviors as a function of the rewarding consequences of the behavior – i.e., the contentment experienced from eating. At each craving tool use ‘ $t$ ’, the expected value for each action ‘ $a$ ’ (eating / not eating) was computed according to Equation 1.

Equation 1):

$$Va(t + 1) = Vat + \alpha(\lambda_t - Vat)$$

Specifically, the expected reward value of an eating behavior at time  $t$  ( $Va_t$ ) was updated for use at the subsequent timepoint ( $Va_{t+1}$ ) using a prediction error term ( $\lambda_t - Va_t$ ). This prediction error reflects the discrepancy between the expected reward value of an eating behavior at time  $t$  ( $Va_t$ ) and the experienced reward at time  $t$  ( $\lambda_t$ ), scaled by a learning rate ( $\alpha$ ) controlling the size of the update at each time point. Here,  $\lambda_t$  at a given craving tool use ( $t$ )

reflects the outcome resulting from the behavior, i.e. the level of contentment experienced after having eaten, and will update steeply if the level of contentment expected ( $Va_t$ ) is highly discrepant from that which was expected. For example, if an individual *anticipated* high levels of reward from eating, yet experienced strong discontentment after *actually* eating the craved food, this would yield a steep decrease in the behavior's reward value due to the discrepancy between the expected and actual outcome resulting from the behavior.

If the participant chose not to eat, the prediction error was coded as 0, and the reward value would remain as that assigned to the previous experience with the behavior  $Va_{(t-1)}$ . At each craving tool use trial, the likelihood of selecting the observed action given the expected values for each action was estimated using the softmax function (Equation 2). The likelihood that a participant would select a given action (eating vs not eating) at each trial was estimated using an inverse temperature parameter (denoted by ' $\beta$ '), controlling the sensitivity with which reward values map onto the probability of selecting an action.

Equation 2):

$$P(\text{eat}(t) | V_{\text{eat}}, V_{\text{no\_eat}}) \\ = \exp(\beta * V_{\text{eat}}(t)) / (\exp(\beta * V_{\text{eat}}(t)) + \exp(\beta * V_{\text{no\_eat}}(t)))$$

The following variables were set as free parameters and estimated for each participant: the learning rate, initial expected eating behavior reward value for each action, and inverse temperature parameter. Free parameters were estimated using an optimization search function which maximized the summed likelihood of observed actions across trials.

Next, in order to compare the goodness of fit of our learning model to our data with that of alternative types of RL models, we also estimated expected values using a learning

model (Rescorla-Wagner / Pearce-Hall Hybrid Model) in which learning rates were dynamically updated at each trial. Thus, at each trial, this model depicts that (Equations 3-4) the learning rate is modulated by an associability term (*assoc*), which is updated based on the prediction error term experienced on the previous trial. For instance, on a trial in which the subject experiences high levels of uncertainty (high prediction error, or high discrepancy between his expected reward and actual reward received), the associability term would update and steeply increase on the next trial and correspondingly modulate the rate at which the prediction error would update the reward value. The associability term has been interpreted as reflecting attention levels, such that highly surprising events might signal to the individual that there is a lot to learn from these experiences and lead to enhanced levels of attention to be paid on the immediately subsequent experiences with the behavior (Le Pelley, 2004). The rate at which the associability term updated at each trial was controlled by an individually fixed parameter (*gamma*). This model is depicted in equations 3 and 4.

Equation 3):  $Va(t + 1) = Vat + assoc(t) * \alpha(\lambda t - Vat)$

Equation 4):  $assoc(t + 1) = gamma * |(\lambda t - Vat)| + (1 - gamma) * assoc(t)$

The likelihood of selecting an action given the reward values assigned to each action were estimated using the softmax rule as described for the Rescorla-Wagner Model. The same free parameters were estimated in the same way as for the previously described Rescorla-Wagner model, with the addition of the individually fixed *gamma* parameter.

## Study2: Methods

### *Data Analysis: Computational Modeling*

For participants in the low-use sample (N = 1044) that had less than 10 craving tool uses, reward values were computed using the same Rescorla-Wagner model as for the high-use sample (Equation 1), and used as variables in multi-level regression models. Parameters for these subjects (learning rate, V0\_no\_eat, V0\_eat, inverse temperature) were estimated using empirical priors obtained from the high-use sample dataset, as has been done in previous work (Gershman, 2016). Here, parameter point estimates for each subject (denoted by  $\hat{\theta}_s$  in Equation 5) were obtained from the maxima of the resulting posterior distributions over parameter values (given the model and data), in which  $p(D_s | \theta_s, m)$  represents the likelihood that the data for subject 's' ( $D_s$ ) were observed given the parameters for this subject (' $\theta_s$ ') and the model (' $m$ '), multiplied by the prior probability distribution for parameters given the model and model hyperparameters ' $\varphi m$ ' (or specific point estimates used as empirical priors). The maximized a posteriori probability estimate for each subject was obtained using optimization procedures within software available at <https://github.com/sjgershm/mfit>.

Equation 5):

$$\hat{\theta}_s = \underset{\theta_s}{\operatorname{argmax}} p(D_s | \theta_s, m) * p(\theta_s | m, \varphi m)$$

Probability distributions of empirical priors consisted of gaussian distributions centered around hyperparameter values (i.e. the mean and standard deviation of free parameters obtained from subjects in the high-use sample). The resulting reward values for each action

(eating vs not eating) were then entered as variables in multi-level regression analyses. In addition, the likelihood that a participant would select a given action (eating vs not eating) at each trial was estimated using the softmax rule computed as in Study 1 and the high-use subsample (Equation 2).

## **Study2: Results**

### *Multi-level regression analyses by High- and Low-Use samples*

The results reported over the aggregated dataset in Study 2 (both high- and low-use subsamples combined) due to their similar pattern are reported here separately for each of these subsamples. Specifically, multi-level regression analyses' effects of TIME (1<sup>st</sup>-level predictor) on reward values (expected and present-moment reward values) and on likelihood of eating estimated from the RW model are reported by High- and Low-Use samples. In addition, multi-level regression analyses' effects of reward values on eating behaviors, and likelihood of eating predictive effect on selected actions (decisions to eat or not) are also reported by High- and Low-Use samples. Finally, a multi-level regression analysis model examining the correspondence between expected and present-moment reward values, using expected reward values as a first-level predictor of present-moment reward values, is reported by High and Low-use samples.

First, multi-level regression analysis revealed a significant negative effect of TIME on expected reward values for both the High-use ( $B = -0.014$ ,  $SE = 0.005$ ,  $t = -3.06$ ,  $p = 0.003$ ) and Low-use sample ( $B = -0.032$ ,  $SE = 0.002$ ,  $t = -20.30$ ,  $p < 0.001$ ), but no significant effect of TIME was observed for present-moment reward values ( $B = -0.002$ ,  $SE = 0.002$ ,  $t = -0.99$ ,  $p = 0.336$  for the High-use sample;  $B = 0.002$ ,  $SE = 0.005$ ,  $t = 0.35$ ,  $p = 0.727$  for the Low-use sample). A significant negative effect of TIME was also observed on likelihood of eating for

the High-use ( $B = -0.013$ ,  $SE = 0.004$ ,  $t = -3.32$ ,  $p = 0.002$ ) as well as Low-use sample ( $B = -0.06$ ,  $SE = 0.003$ ,  $t = -17.82$ ,  $p < 0.001$ ).

In addition, multi-level regression analyses revealed a significant positive predictive effect of expected reward values on eating frequency ( $B = 0.25$ ,  $SE = 0.042$ ,  $t = 6.04$ ,  $p < 0.001$  for the High-use sample;  $B = 1.40$ ,  $SE = 0.053$ ,  $t = 26.33$ ,  $p < 0.001$  for the Low-use sample) and eating intake ( $B = 0.51$ ,  $SE = 0.100$ ,  $t = 5.19$ ,  $p < 0.001$  for the High-use sample ;  $B = 2.91$ ,  $SE = 0.137$ ,  $t = 21.34$ ,  $p < 0.001$  for the Low-use sample) for each High and Low-use samples. With respect to present-moment reward values, this variable also predicted eating frequency ( $B = 0.51$ ,  $SE = 0.041$ ,  $t = 12.50$ ,  $p < 0.001$  for the High-use sample ;  $B = 0.46$ ,  $SE = 0.021$ ,  $t = 22.04$ ,  $p < 0.001$  for the Low-use sample) and intake ( $B = 1.10$ ,  $SE = 0.100$ ,  $t = 10.99$ ,  $p < 0.001$  for the High-use sample;  $B = 1.06$ ,  $SE = 0.049$ ,  $t = 21.33$ ,  $p < 0.001$  for the Low-use sample) for each sample. Likelihood of eating also significantly predicted decisions to eat or not for the High ( $B = 1.02$ ,  $SE = 0.044$ ,  $t = 23.14$ ,  $p < 0.001$ ) and Low-use sample ( $B = 1.06$ ,  $SE = 0.017$ ,  $t = 62.38$ ,  $p < 0.001$ ).

Finally, analyses revealed that expected reward values significantly predicted present-moment reward values in each High- ( $B = 0.12$ ,  $SE = 0.04$ ,  $t = 2.83$ ,  $p = 0.007$ ) and Low-use samples ( $B = 0.32$ ,  $SE = 0.05$ ,  $t = 6.86$ ,  $p < 0.001$ ).

## References

- Epel, E. S., Tomiyama, A. J., Mason, A. E., Laraia, B. A., Hartman, W., Ready, K., Acree, M., Adam, T. C., St Jeor, S., & Kessler, D. (2014). The reward-based eating drive scale: a self-report index of reward-based eating. *PLoS One*, *9*, e101350. <https://doi.org/10.1371/journal.pone.0101350>.
- Gershman, S. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology*, *71*, 1–6. <http://doi.org/10.1016/j.jmp.2016.01.006>.
- LePelley, M. E., & McLaren, I. P. (2004). Associative history affects the associative change undergone by both presented and absent cues in human causal learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(1), 67–73. <https://doi.org/10.1037/00097-7403.30.1.67>
- Meule, A., Hermann, T., & Kubler, A. (2014). A short version of the Food Cravings Questionnaire-Trait: the FCQ-T-reduced. *Frontiers in Psychology*, *5*, 190. <https://doi.org/10.3389/fpsyg.2014.00190>.
- Meule, A., Reichenberger, J., & Blechert, J. (2018). Development and preliminary validation of the Salzburg Stress Eating Scale. *Appetite*, *120*, 442–448. <https://doi.org/10.1016/j.appet.2017.10.003>.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Williams, E. P., Mesidor, M., Winters, K., Dubbert, P. M., & Wyatt, S. B. (2015). Overweight and obesity: Prevalence, consequences, and causes of a growing public health problem. *Current Obesity Reports*, *4*(3), 363–370. <https://doi.org/10.1007/s13679-015-0169-4>.