**Supplementary Materials:**

**METHODS:**

**Sample size calculation:**

Sample size was calculated on the basis of a minimally meaningful effect size of *f =0.2* in reducing chocolate cue reactivity and overconsumption using G\*Power 3.1.9. For 90% power in a mixed 3 (Time: baseline, post-manipulation, follow-up) x 3 (Group) ANOVA, with α =.05 and assumed repeated measures correlation of *ρ* =0.5 a N of 69 (23 per group) was required. Expecting minimal attrition, we therefore aimed to recruit 75 total participants (25 per group).

**Statistical Approach**

Data Handling: Data analysis was performed using IBM SPSS 25 and *R* for Windows. Primary in-lab variables (cue reactivity, attentional bias, progressive ratio task and state craving) were assessed using mixed ANOVA with a within-subjects factors of cue type (HPF vs. LPF) and *Time* (*Baseline* vs. *post-manipulation*). For questionnaire measures of chocolate craving and disordered food consumption (FCQ-T, ACQ, BES, TFEQ, RS food diary chocolate consumption and TLFB), the *Time* factor had three levels (*Baseline* *post-manipulation, follow-up*). All analyses included a between-subjects factor of Group (*RET+PBO*, *RET+RAP*, *No RET + RAP*). Pearson’s correlations were performed between prediction error ratings and key outcomes. For the above analyses, if this correlation was significant it was included as a covariate in the models above model. For all analyses, the specified models were fully factorial. If a main effect of *Group* was observed (in the absence of higher-level interactions), this was followed up with pairwise comparisons across groups. If a *Group x Time* interaction is observed, this was followed up by assessing the simple effect of *Time* within each *Group*.

All data were assessed for outlying data points; defined as those > +/- 3 standard deviations from the group mean. These data points were Winsorized according to the recommendations of Tabachnick and Fidell 45. Screened data were then assessed for dispersion with visual inspection of histograms and Q-Q plots. Where sphericity was violated in repeated measures, the Greenhouse-Geisser correction or multivariate terms were used, depending on ε values and according to the recommendations of Stevens47. Significant k >2 main effects and interactions in omnibus ANOVAs were investigated with multivariate simple effects analyses and paired tests on marginal means, where appropriate. Alpha for all *a priori* tests was 0.05(2-sided), with *p*-values Sidak- corrected for post-hoc tests. False discovery rate in analysis of baseline demographic variables was controlled with the Benjamini-Hochberg procedure. All analyses of primary outcomes were performed blind by RKD and the blinding code not broken until analysis was completed. The pre-registered analysis plan can be found on the Open Science Framework (<https://osf.io/tqxdb> DOI 10.17605/OSF.IO/TQXDB).

**Tasks**

**Oculomotor Bias**

This visual probe task assessed attentional capture by chocolate images by pairing with non-binge food (LPF) images. Image pairs were presented side-by side on screen and aye movement to the image assessed. The primary eye-tracking measures were summed fixation on each image in each trial (*Dwell time*), latency to first fixation on each image from trial start (*fixation latency*) and duration of this first fixation. Full task details are given in the supplementary material.

This task tracked eye movements using a Tobii T-60 eye-tracker (Tobii, Sweden), to assess oculomotor attentional capture by chocolate cue images an index of the salience of these images. The stimuli were eight HPF chocolate images and eight LPF images from the *FoodPics* database. Images were presented in pairs, side-by-side on screen on each trial, separated by a distance of 400 pixels. After 2000ms, the image pairs disappeared and a ‘probe dot’ appeared in the previous location of the HPF/chocolate image (congruent) or LPF image (incongruent). Participants responded as to the location of the probe dot as quickly and accurately as possible using the keyboard. Each HPF image was paired twice with each of the LPF images, counterbalanced for presentation side of the chocolate image (left/right) and congruence of the probe dot (congruent/incongruent), yielding 128 total trials. The task was implemented in PyGaze35. For the visual probe task, following Mogg et al 46, first fixation latencies below 100ms were trimmed from average first fixation calculation for that participant, as this represents non-attentional orienting.

**Motivation to consume chocolate: Progressive Ratio Task.**

On each trial, participants first chose between three options: 1) performing a timed tapping task to win a standard amount of chocolate (3 grams: one Cadbury’s milk chocolate button, Bourneville, UK) 2) The same task for a non-binge food (one 3g dried strawberry slice), or 3) ending the task. They made this choice by clicking on an image of chocolate or strawberry or an ‘end task’ button. After selecting a food option, participants were required to tap the space bar a set number of times within a limited time period in order to ‘earn’ the selected food. The time limit was adjusted on a trial-by trial basis such that roughly 3 presses/second were required. Each time food was successfully ‘earned’, participants were given up to one minute to consume the earned food. They had to consume the food before continuing the task. After each ‘successful’ food trial, the number of taps required to earn that food on the next trial increased by twenty. There was no fixed number of trials, as the task ended when the participant selected the option to end on the choice screen. After each successful trial, participants rated their enjoyment of the food consumed from -5 (extremely unpleasant) to +5(extremely pleasant) and their subjective hunger from -5 (completely full) to +5 (extremely hungry). The primary extracted indices were 1) number of choices for chocolate vs. strawberries 2) The ‘break point’ in number of required taps for the last trial participants decide to play for a food type 3) An action-incentivisation index for each cue type calculated as $\frac{1}{mean RT } ×N choices$ (where mean RT = mean reaction time per press), which could account for the lack of motivation to consume where no choices for a particular food type were made.

**Chocolate consumption diary**: An online diary was used to obtain data on levels of naturalistic chocolate consumption in the week preceding (*baseline*) and following (*post-manipulation*) manipulation and at one month post-*Day 1* (*follow-up*). On each day, participants received a text prompt to record the chocolate they consumed that day, their *peak level* of chocolate craving, their subjective opinion of whether they overconsumed and binged on chocolate or any other food. Chocolate consumption was converted to a grams equivalent, peak craving was on a 0-100 visual analogue scale and overconsumption/ binge occurrence was a daily binary response. On *Day 1* and *Day 10*, a Timeline Follow-Back calendar-based measure of chocolate consumption (in grams) was used to ensure consumption data were available for the key peri-manipulation period.

**Drug preparation:**

Drugs were stored in numerically coded envelopes. Although closely matched in appearance, due to non-identical formulations of drug and placebo, tablets were administered by a secondary researcher not involved in the study and participants were blindfolded while taking the tablets to maintain the double-blind. All drugs were given 60 minutes prior to the RET/No RET procedure to coincide central rapamycin availability with the post-retrieval ‘reconsolidation window’.

**RESULTS**

**Naturalistic Chocolate and Food Consumption: Timeline follow-back.**

Timeline follow-back logged frequency of chocolate consumption (% of days) showed a main effect of *Day* [F(2, 138) =14.714, *p<* .001, *η2p*=.176], with chocolate consumption reducing between *Day 1* and *Day 10* [t(71)=3.341, *p=*.004, *r* =.369] and remaining stable from *Day 10* to follow-up [t(71)=1.91, *p*=.17, *r* =.221]. No differential effects of *Group* were observed.

**Manipulation check: Peri-Retrieval Measures**

Comparison of cues rated during the RET and No RET procedures showed significant differences between groups in displayed food *liking* [F(2, 70) =12.502, *p<* .001, *η2* =.263], *wanting* [F(2, 70) =9.322, *p<* .001, *η2* =.21] and *binge risk* [F(2, 70) =32.513, *p<* .001, *η2* =.482]. In all cases, this represented lower ratings of LPF images in the *No RET* group than ratings of chocolate images in *RET + RAP* and *RET+PBO* (all *p*s < .001), with no difference between ratings in the two *RET* groups (all *p*s ≥. 258). The same pattern of responding was observed for participants’ ratings of the food UCS itself. This indicates that the retrieval procedures were effective in differentially activating consumption-related motivational memory processes, as intended.

Surprise (prediction error) ratings also differed significantly between groups [F(2,71) =38.512, *p<*.001, *η2* =.52], with significantly higher surprise in *RET+PBO* and *RET + RAP* than *No RET + RAP* [*t*(47)=8.288, *p*<0.001, *r*=.771] and *t*(47)=6.712, *p*<0.01, *r*=.7, respectively] and no difference in surprise between the two *RET* groups [*t*(49)=1.592, *p*=.309, *r* =.222]. This is commensurate with the withholding of the food UCS (chocolate) in the *RET* groups and consumption of the food, as expected in *No RET*. The procedure was therefore successful in generating negative prediction error in *RET* groups.

The RET groups did not differ in rated liking of, urge to consume or binge risk on cue-depicted foods (chocolate) nor on the chocolate UCS itself, during retrieval. There was a non-significant trend for greater rated surprise in *RET+PBO* than *RET+RAP* [F(1,48) =3.641. *p=*.062. *η2* =.07].

**Drug side effects and drug guess:** Chi square analysis of drug condition (placebo, rapamycin) x Side effect experience during drug session (yes/no) was non-significant (χ22=0.951, *p*=0.330), suggesting participants who received rapamycin did not report a greater number of side effects than those who received placebo. Chi square analysis of Group x Drug guess found a significant effect of group (χ2 (2)=7.721, *p*=0.021), with participants in *RET+RAP* guessing they received rapamycin with greater frequency than the other two groups. Drug guess Ns per group were: *RET+PBO* = 4; *No RET+RAP*=7; *RET+RAP*=13.