

Replication of Boswell et al 2018

“Training in cognitive strategies reduces eating and improves food choice”

PNAS 2018 115(48), E11238-E11247.

<https://www.pnas.org/content/115/48/E11238>

The original paper includes several studies. We randomly chose study 4. In this between-subject experiment, participants start by making hypothetical choices between images of food pairs. Thereafter, participants are randomized to either the Regulation of Craving Training with negative strategies (negative ROC-T) treatment, the Regulation of Craving Training with positive strategies (positive ROC-T) treatment, or to receive no training and instead look at images of food items (control). Participants thereafter again make hypothetical choices between images of food pairs. We focus on the comparison of the negative ROC-T and the control condition. There is a higher increase in healthy hypothetical food choices for participants in the negative ROC-T group compared to the control group.

Hypothesis to replicate and bet on: Participants who receive training in Regulation of Craving (ROC-T) with negative strategies subsequently increase healthy food choices in the presence of tempting alternatives. The authors test the above hypothesis in an independent samples *t*-test comparing the participants’ change (posttest – pretest) in hypothetical healthy food choices between the negative ROC-T and the control condition. (independent samples *t*-test (not assuming equal variances): negative ROC-T vs. control: $t(173.36) = 5.37, p < 0.001$); p. 12 in the SI.

Criteria for replication: The criteria for replication are an effect in the same direction as the original study and a *p*-value < 0.05 in a two-sided independent samples *t*-test.

Power analysis: The original study had 260 participants after exclusions (106 in the negative ROC-T treatment and 154 in the control treatment). The standardized effect size (Cohen’s *d*) was $d = 0.666$. To have 90% power to detect 67% of the original effect size, a sample size of $n = 213$ is required. Since we require the replication sample size to be at least as large as in the original study, the replication experiment will use a sample size of $n = 260$ and the power will thus be $>90\%$ to detect 67% of the original effect size.

Sample: Only participants from the US were allowed to participate in the original study, and previous participation in any study by the original authors’ group precluded participation in subsequent studies. Across all studies, 10.9% of participants were excluded for incomplete data reporting, and 12.7% of participants were excluded for failing manipulation checks. The following exclusion criteria were used in the original study: comprehension check filters (two independent coders flag free response answers for non-comprehension of the essay prompt. Differences are reconciled through dialogue), bot check filters (described in Appendix p. 1), multiple submissions, not reading all essay screens, not completing the choice task, not completing the ROC task, completing less than half (36) of the ROC trials, reading essay screens for less than 10 seconds, and skipping more than 54 choices. Due to expected manipulation check failures and incomplete data reporting we will overrecruit participants to ensure to reach the required sample of 260. We will apply the same exclusion criteria, we will make sure that participants can only participate once from the same account in this specific study, we will exclude participants that participated in any study by the original authors’ group (with the help of the original authors) and we will only recruit participants with a HIT approval rate of 95% or higher. We will also check all IP addresses via <https://www.ipqualityscore.com/>; and we will remove any participants where one or more of the following is true: fraud score \geq

85; TOR = True; VPN = True; Bot = True; abuse velocity = high. The replication sample size is the sample size after any exclusions of participants.

Materials: We will use the same material as in the original study, kindly provided by the original authors. In particular, the experiment will be conducted using the original *Inquisit* survey, which was adapted by the original authors to only run the negative ROC-T treatment and the control treatment. The experiment uses the same titles and photographs.

Procedure: We will closely follow the procedure of the original experiment. We will only replicate the conditions negative ROC-T and control. In the original study another treatment featuring positive ROC-T was included which we will not replicate. Excluding this treatment does not affect the relevant data collection. The following summary of the experimental procedure is therefore largely based on the description of the experiment in the article (p. E11240) and the section “Methods” (pp. E11245-E11246).

Participants will first be shown a Captcha, and will thereafter provide informed consent. After this we will include an attention check that participants will need to pass to continue to the study. This attention check is in addition to any other potential attention check(s) used in the original study. In a between subject design, participants will be randomized to either the negative ROC-T treatment or the control. Participants will conduct three tasks: (i) the Choice Task, (ii) ROC-T, and (iii) the Choice Task for a second time. During each Choice Task, participants will be asked to make hypothetical choices where they pairwise choose the food item they desire more in 108 pairs of food (where 72 key choices come from healthy-unhealthy food pairings; the 108 pairs are the same during both Choice Tasks). During the ROC-training participants will either receive the negative ROC training or no-training (control). In the negative ROC-training, participants will read a brief essay about the risks of eating unhealthy food, complete six free-response questions about the content of the essays and are then trained to inform a cognitive strategy so that they frame unhealthy foods in terms of negative qualities when they rate their craving and willingness to pay. In the control condition, participants will view images and rate craving and willingness to pay for each item. This phase will last ~15 min.

Analysis: The analysis will be performed as in the original paper. The analysis code was kindly provided by the original authors. In particular, we will perform an independent samples *t*-test (not assuming equal variances) of differences in the choice of healthy foods after the intervention minus choice of healthy foods before the intervention divided by total number of relevant choices (which is 72) between subjects in the negative ROC-T condition and subjects in the control condition. As advised by the original authors, we will also perform Levene’s test of equal variances. If the Levene’s test is non-significant, we will additionally report the *t*-test above assuming equal variances. The conclusion about whether the study replicates or not will only be based on the main replication test (i.e., the *t*-test assuming unequal variances). The *t*-test assuming equal variances are mainly relevant for understanding why the study failed to replicate if it should fail to replicate.)

Subject payments: We are standardizing payments across all replications so that studies have a certain show-up fee depending on the expected length of the study, with an hourly wage from the show-up fee of \$8 and a minimum payment of \$1 (for studies with incentive payment we use the same incentive payment as in the original study; and this payment is paid in addition to the show-up fee). If we have problems recruiting, we will increase the show-up fee.