Do impulsive individuals benefit more from food go/no-go training? Testing the role of inhibition capacity in the no-go devaluation effect

4 authors:

Zhang Chen
Radboud University
7 PUBLICATIONS   **59** CITATIONS

SEE PROFILE

Harm Veling
Radboud University
58 PUBLICATIONS   **1,092** CITATIONS

SEE PROFILE

Ap Dijksterhuis
Radboud University
161 PUBLICATIONS   **12,321** CITATIONS

SEE PROFILE

Rob W Holland
Radboud University
95 PUBLICATIONS   **3,405** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Changing Food Choices With the Go/No-Go Training View project

Do Impulsive Individuals Benefit More From Food Inhibition Training?

Testing the Role of Inhibition Capacity in the No-Go Devaluation Effect

Zhang Chen, Harm Veling and Ap Dijksterhuis

Radboud University, Nijmegen, the Netherlands

Rob W. Holland

Radboud University, Nijmegen, the Netherlands

University of Amsterdam, Amsterdam, the Netherlands

Author Notes

Zhang Chen, Behavioural Science Institute, Radboud University; Harm Veling, Behavioural Science Institute, Radboud University; Ap Dijksterhuis, Behavioural Science Institute, Radboud University; Rob W. Holland, Behavioural Science Institute, Radboud University and Faculty of Social and Behavioural Sciences, University of Amsterdam.

Correspondence concerning this article should be addressed to Zhang Chen, Montessorilaan 3, A.09.02, Radboud University, 6500 HE, Nijmegen, the Netherlands. E-mail: z.chen@psych.ru.nl

Abstract

Recent research showed not responding to food items in a go/no-go task leads to devaluation of these food items, which may help people regulate their eating behaviors. The Behavior Stimulus Interaction (BSI) theory explains this devaluation effect by assuming that inhibiting impulses triggered by appetitive foods elicits negative affect, which in turn devalues the food items. BSI theory further predicts that the devaluation effect will be stronger when food items are more appetitive and when individuals have low inhibition capacity. To directly test these hypotheses, in the current study we manipulated the appetitiveness of food items and measured individual inhibition capacity with the stop-signal task. Food items were consistently paired with either the go or no-go cue during training, so that participants responded to some items and not to others. Evaluations of these food items were measured before and after the training. Across two preregistered experiments, we consistently found after the training no-go foods were liked less compared to both go foods and foods not used in the training. Unexpectedly, devaluation effect occurred for both appetitive and less appetitive food items, and exploratory signal detection analysis suggested this might be explained by comparable learning of the contingencies for these items. Furthermore, the strength of devaluation did not consistently correlate with individual inhibition capacity, and Bayesian analyses combining data from both experiments provided moderate support for the null hypothesis. In sum, the current project reliably demonstrated the devaluation effect induced by the go/no-go training, but failed to obtain evidence for BSI theory. While more work is still needed, the result reinforces the notion that the go/no-go training may be a promising tool to help people regulate their eating behaviors.

Keywords: food, go/no-go task, devaluation, response inhibition, stop-signal paradigm, impulsivity

Do Impulsive Individuals Benefit More From Food Inhibition Training?

Testing the Role of Inhibition Capacity in the No-Go Devaluation Effect

Today's world is filled with appetitive foods: in kiosks at train stations, in supermarkets and grocery stores, on television, appetitive foods are everywhere. Appetitive foods automatically attract our attention (Nijs, Muris, Euser, & Franken, 2010), activate the brain's reward region (Wang et al., 2004), and excite the motor system (Gupta & Aron, 2011), so that we may easily notice, desire, and eventually obtain and consume them. Because of our innate preference for calories (Breslin, 2013), appetitive foods are often energy-dense and contain much sugar and fat. Excessive intake of appetitive yet energy-dense foods is broadly considered a major contributor to increased body weight and the worldwide obesity epidemic we are facing right now (World Health Organization, 2016).

Although many people are living in an obesogenic environment, gaining weight is not inevitable. Appetitive energy-dense foods may trigger potent impulses and urges, but people also have the capacity to inhibit these impulses and regulate their thoughts and behaviors. This capacity to inhibit predominant responses is termed *inhibition*, which is one of three main executive functions (*inhibition*, *updating* and *shifting*, respectively; Miyake & Friedman, 2012). Individuals differ in their capacity to inhibit impulses, and this difference in inhibition capacity may have implications for regulating eating behaviors. For instance, individuals with relatively low inhibition capacity consume more appetitive foods in the lab (Guerrieri et al., 2007), and their consumption of foods is more strongly determined by their automatic affective reactions to foods than individuals with high inhibition capacity (Hofmann, Friese, & Roefs, 2009). People with low inhibition capacity and strong implicit preferences for snack foods also gain most weight over a year (Nederkoorn, Houben, Hofmann, Roefs, & Jansen, 2010). Furthermore, obese

and overweight individuals show lower inhibition capacity in comparison to healthy controls (Kulendran et al., 2014; Nederkoorn, Braet, Van Eijs, Tanghe, & Jansen, 2006a; Nederkoorn, Smulders, Havermans, Roefs, & Jansen, 2006b; for a meta-analysis, see Lavagnino, Arnone, Cao, Soares, & Selvaraj, 2016), suggesting that low inhibition capacity may be a contributing factor to overweight.

If the problem in regulating eating behaviors stems from the difficulty in inhibiting impulses evoked by appetitive foods, reducing the strength of the impulses triggered by foods should help people to better regulate their eating behaviors. Inspired by this idea, several training procedures have been developed, and one such training is the go/no-go training (GNG training), in which participants consistently respond to some stimuli (e.g., by pressing a key, go trials) and withhold their responses to other stimuli (e.g., do not press any key, no-go trials). When used as training on foods, appetitive energy-dense foods can be consistently presented on no-go trials so that participants do not respond to these foods. Previous research showed that after GNG training, appetitive foods are evaluated less positively (Veling, Aarts, & Stroebe, 2013a; Chen, Veling, Dijksterhuis & Holland, 2016) and are chosen less often (Veling et al., 2013a; Veling, Aarts, & Stroebe, 2013b). When offered a chance to consume these appetitive foods in the laboratory, participants consume fewer of them if these foods have been paired with no-go trials in the training (Houben, 2011; Houben & Jansen, 2011; Houben & Jansen, 2015; Lawrence, Verbruggen, Morrison, Adams, & Chambers, 2015b; Veling, Aarts, & Papies, 2011; Folkvord, Veling & Hoeken, 2016). Moreover, repeated training with high-calorie foods has been shown to facilitate weight loss attempts in two studies (Lawrence et al., 2015a; Veling, van Koningsbruggen, Aarts, & Stroebe, 2014; for a recent meta-analysis, see Allom, Mullan, & Hagger, 2015).

Not responding to appetitive foods thus leads to decreased evaluation. One explanation for such devaluation effect is offered by the Behaviour Stimulus Interaction theory (BSI theory, Veling, Holland, & van Knippenberg, 2008). According to the BSI theory, exposure to appetitive foods triggers an impulse to approach the foods. However, when these foods are presented together with a no-go cue, participants need to engage in response inhibition in order to overcome this impulse. This conflict between the impulse to respond and response inhibition elicits negative affect. After repeated pairings, the negativity of the conflict is attached to the appetitive foods, leading to decreased evaluation of them.

The BSI theory posits that the strength of the conflict is jointly determined by two competing processes: the initial impulses triggered by the food items, and the response inhibition process to overcome and inhibit these impulses. The first prediction from the BSI theory is that the devaluation effect induced by response inhibition in GNG training should be stronger for stimuli that are more appetitive, since appetitive stimuli trigger stronger impulses, and the strength of the resulting conflict is accordingly higher. In line with this idea, previous research has shown that GNG training is more effective in lowering the evaluations of positive stimuli than that of neutral or negative stimuli (Veling et al., 2008; Chen et al., 2016; but see Frischen, Ferrey, Burt, Pistchik, & Fenske, 2012). The influence of GNG training on food choices is also stronger when people have a relatively high appetite, presumably because individuals with a high appetite experience strong impulses toward appetitive foods (Veling et al., 2013a). Furthermore, several studies showed that GNG training is especially effective for restrained eaters (Houben, 2011; Houben & Jansen, 2011; Lawrence et al., 2015b; Veling et al., 2011). Restrained eaters chronically restrict food intake to lose weight, but are mostly unsuccessful and often end up consuming more than non-restrained eaters (Fedoroff, Polivy, & Herman, 1997; Stirling &

Yeomans, 2004). One reason for their failure to control body weight is that restrained eaters respond more strongly to appetitive foods (Brunstrom, Yates, & Witcomb, 2004; Harvey, Kemps, & Tiggemann, 2005; Papies, Stroebe, & Aarts, 2007; Veenstra & de Jong, 2010), and this might again explain why GNG training is more effective for them.

The second prediction from the BSI theory is that the effectiveness of GNG training depends on the response inhibition process. In line with this idea, recent research found that increasing the proportion of no-go trials in GNG training, thereby presumably reducing the engagement and activation of response inhibition on no-go trials, renders the GNG training procedure unsuccessful in triggering devaluation, suggesting a causal role of response inhibition in inducing devaluation (Chen et al., 2016; for discussions on to what degree GNG training evokes motor inhibition toward the stimuli, also see Wessel, O'Doherty, Berkebile, Linderman, & Aron, 2014; Veling, Lawrence, Chen, van Koningsbruggen, & Holland, 2017). Furthermore, Adams and colleagues (2017) recently showed that inhibition trainings (GNG training and stop-signal training) with higher rates of inhibition accuracy are more effective in modifying food intake, corroborating the idea that successful response inhibition leads to devaluation. Another way to address this question is by examining individual differences in inhibition capacity. More specifically, for individuals with low inhibition capacity, inhibiting an impulse might be more difficult and conflict-inducing than for individuals with high inhibition capacity. The devaluation effect may therefore be larger when individuals have lower inhibition capacity. Indirect evidence came from studies with restrained eaters (Houben, 2011; Houben & Jansen, 2011; Lawrence et al., 2015b; Veling et al., 2011). As mentioned above, GNG training is more effective for restrained eaters, and as restrained eaters in general have lower inhibition capacity (Nederkoorn, Van Eijs, & Jansen, 2004; measured by the stop-signal task, Logan, Cowan, & Davis, 1984), this

pattern of results could thus be interpreted as in line with this second prediction. Note that since restrained eaters are both highly responsive to food cues and have low inhibition capacity, these findings cannot disentangle the two predictions made by the BSI theory. More direct evidence came from a study by Houben (2011), in which participants first received the stop-signal task to measure their inhibition capacity, and then performed response inhibition training (albeit a training based on the stop-signal task, not the GNG training) on high-calorie foods. Food intake was reduced after the training, but only for individuals with low inhibition capacity, indicating that the training is indeed more effective for individuals with low inhibition capacity.

Combined, these two predictions suggest an interesting and useful feature of the GNG training. That is, GNG training becomes more effective when food items are more appetitive and when people have low inhibition capacity to inhibit their impulses – the situation predicting excessive intake of calories (Nederkoorn et al., 2010). Although all the findings reviewed above are in line with these two predictions, the evidence is still relatively scarce, especially for the role of inhibition capacity in such devaluation effect. Furthermore, to the best of our knowledge, these two predictions have never been jointly tested in one design. This interactive process between the impulsive approach tendencies and the response inhibition process, as outlined by the BSI theory, is therefore still not entirely clear. Further examining this process will not only provide more theoretical insight into the underlying mechanism of the devaluation effect, but will also be useful for applied use. For instance, knowing which individuals (e.g., low on inhibition capacity) may benefit more from GNG training in which situations (e.g., for food items for which they have strong cravings) could enable us to better tailor and personalize GNG training to maximize the efficiency of the training.

In the current research, we used evaluations of foods as the dependent measurement, since the liking for foods has been considered the strongest motivation underlying people's food choices (Phan & Chambers, 2016) and has been shown to mediate the effect of GNG training on food choice (Veling et al., 2013a). The change in evaluation also correlates with change in body weight after GNG training (although the mediation effect was not significant; Lawrence et al., 2015a). We included food items that were evaluated as highly appetitive and food items that were evaluated as less appetitive, to test the role of approach tendencies in devaluation, as more appetitive foods are assumed to trigger stronger approach tendencies (Seibt, Häfner, & Deutsch, 2007). To measure individual inhibition capacity, we used the stop-signal task, because this task allows the estimation of the covert latency of stopping and is currently one of the most widely used measurements of response inhibition (Verbruggen & Logan, 2008).

Experiment 1

To reduce flexibility in data analysis, and to be able to distinguish clearly between confirmatory and exploratory analyses, we preregistered the sample size, analyses plans and predictions for this experiment. More specifically, we predicted that high-rated no-go foods would be devalued in comparison to both go and untrained foods (i.e., foods not used in GNG training, see method below), while no devaluation or a smaller devaluation effect was predicted for low-rated no-go foods. Furthermore, we expected that lower inhibition capacity would be associated with stronger devaluation effect for high-rated foods, while no specific association was predicted for low-rated foods. The preregistration can be found at osf.io/cz8zr/.

Method

Sample Size. Based on our previous work on stimulus evaluation in which a similar task was used (Chen et al., 2016), 30 participants were required to achieve 80% power for the

devaluation effect (defined as the difference between no-go items and items not used in the training. Using the difference between no-go and go items gave a smaller sample size in power analysis). Note that in previous work, participants were often required to fast for a few hours before the experiment (e.g., Chen et al., 2016; Adams et al., 2017), to control the level of appetite. However, such requirement to fast might increase their overall appetite and make all food items appear appetitive (Veling et al., 2013a), reducing the difference between highly appetitive and less appetitive food items. In the current study, we therefore did not ask participants to fast before the experiment. Furthermore, both high- and low-rated items were used, which led to fewer stimuli in each cell. To compensate for the potential decrease in effect size due to these two differences, we planned to recruit 70 participants. Sensitivity analysis revealed that with 70 participants, we were able to detect a devaluation effect as small as Cohen's $d_z = 0.340$ with 80% power. For the correlation between inhibition capacity and no-go devaluation effect, no prior effect size was found. Sensitivity analysis showed that with 70 participants, a correlation effect no smaller than .327 can be detected with 80% power.

Participants. Participants, mainly students at Radboud University, were recruited via the online participation system. They received either course credit or money as compensation for their participation.

Materials. Eighty pictures of various foods (e.g., desserts, fruits, vegetables, candies, full meals etc.) were selected from the food-pics database (Blechert, Meule, Busch, & Ohla, 2014; for the images used, see osf.io/cz8zr/). The procedure was implemented in PsychoPy (version v1.81.03; Peirce, 2007) and run on a Windows 7 computer individually for each participant.

Procedure.

Stop-Signal Task. Upon arrival, participants first received a stop-signal task to measure their inhibition capacity. A circle was presented on the screen, and for each trial, an arrow pointing either to the left or to the right was presented in the middle of the circle. Participants' primary task was to indicate the direction of the arrow by pressing F or J, for left or right respectively. Each arrow was presented for 1000 ms and the inter-trial interval randomly varied between 1000 ms and 2000 ms, in steps of 100 ms. Occasionally (on 25% of the trials) a beep (i.e., the stop signal, 700 Hz, 300 ms) was played via a headphone, signaling to the participants that they should not respond for the current arrow. This stop signal was played at a delay (i.e., the stop-signal delay, SSD) from the arrow onset. SSD was initially set to 250 ms and dynamically adjusted based on participants' performance: SSD increased by 50 ms after a successful stop, and decreased by 50 ms after a failed stop. This tracking procedure ensures the overall probability of stopping when a stop signal occurs would be around 50%. The whole stop-signal task contains 4 blocks, with 20 stop-signal trials and 60 no signal trials in each block. After each block, participants took a short break, and the performance in the current block (e.g., average goRT, p(stop|signal) and average SSD) was provided to the experimenter. Based on the performance, the experimenter instructed participants to either respond faster or to slow down (Verbruggen et al., 2013). The behavioral parameters we aimed to achieve were: goRT between 400 ms and 650 ms, p(stop|signal) between .4 and .6, and the average SSD above 100 ms (see Wessel et al., 2014, in which the same behavioral parameters were adopted).

Unrelated Questionnaires. After finishing the stop-signal task, participants filled out a few questionnaires on self-forgiveness, rumination and mindfulness for another study. Details and data from these questionnaires will not be further discussed.

Rating Before the Training. Participants then received a self-paced evaluation task, in which they rated each of the 80 food images on how attractive each food looked to them at the moment, by using a 200-point slider (from -*100 = not attractive at all* to *100 = very attractive*).

Ranking and Selection. The experimental program then rank ordered the 80 food items from the highest rating to the lowest. Twenty-five high-rated (ranked 11-35) and 25 low-rated (ranked 46-70) items were selected. For both high-rated and low-rated foods, the 25 images were further divided into 5 sets, with the average evaluation for each set matched.

Go/No-Go Training. Of the 5 sets of high-rated foods, 3 sets were randomly selected as go stimuli, 1 set selected as no-go stimuli and the remaining set was not used in the training and served as an untrained baseline. The same selection was used for the 5 sets of low-rated foods. During the GNG training, a food image was presented in the middle of the screen. 100 ms after image onset, a beep was played via headphone (1000 Hz, 300 ms or 400 Hz, 300 ms; the assignment of beeps as go and no-go cues was counterbalanced across participants; see Chen et al., 2016). If they heard a go cue, participants were instructed to press the B key on the keyboard as fast as possible, before the image disappeared from screen; if they heard a no-go cue, they were instructed to withhold their response and just wait till the image disappeared. In either case, the image stayed on screen for 1000 ms. No performance feedback was provided in the experimental blocks, and the inter-trial interval randomly varied from 1000 to 1500 ms, in steps of 100 ms. The training started with a practice block of 10 trials (different images were used in practice), and then proceeded with 6 experimental blocks. Each of the 40 images was randomly presented once in each block, and the whole training consisted of 240 trials.

Rating After the Training. After the training, participants received a rating task that was the same to the one they received before the training. Only the 50 selected images were used in this second rating task.

Memory Recall. After that, participants received a memory recall task, in which each of the 40 images that were shown in the GNG training was presented again. For each image, they had to indicate whether it was associated with pressing B or not pressing B. The memory recall task was included to assess participants' learning of the stimulus-response associations.

Demographics. For exploratory reasons, participants filled out the restraint eating scale (Herman & Polivy, 1980), reported whether they were currently on a diet, the last time they consumed food, whether they had lunch or not, their current hunger level, weight and height.

Results

Participants. Seventy-one participants (8 males, 63 females, $M_{age} = 20.7$ years, $SD = 2.74$) took part in the experiment. No participant met the preregistered exclusion criteria (i.e., accuracy on go or no-go trials 3SDs below sample mean and below 90%) and all data were included in the analyses.

Performance in GNG Training. The average accuracy on go trials was 99.0%, $SD = 1.48$%; the average accuracy on no-go trials was 89.6%, $SD = 9.67$%. For the calculation of mean go reaction time (go RT), the incorrect go trials were first removed. The go trials on which the RT was 3SDs from a participant's own mean were further removed and the average go RT on the remaining go trials was then calculated for each participant. The average go RT was 397.8 ms, $SD = 45.7$. Overall, participants performed well in the GNG training. For the performance across blocks in the training, see the supplementary material.

Stimulus Evaluation. A repeated-measures ANOVA (rmANOVA) with rating level (high vs. low), measurement time (pre- vs. post-training) and training condition (go vs. no-go vs. untrained) as within-subject factors and the average rating as dependent variable was conducted. The Greenhouse-Geisser correction was used when the assumption of sphericity was violated. The results showed that the three main effects were all statistically significant: for the main effect of rating level, $F(1, 70) = 488.7$, $p < .001$, $\eta_p^2 = .875$; for the main effect of measurement time, $F(1, 70) = 25.4$, $p < .001$, $\eta_p^2 = .266$; for the main effect of training condition, $F(1.69, 118.1) = 15.4$, $p < .001$, $\eta_p^2 = .181$. The interaction between rating level and measurement time was significant, $F(1, 70) = 48.8$, $p < .001$, $\eta_p^2 = .411$; the interaction between measurement time and training condition was also significant, $F(1.70, 119.1) = 14.9$, $p < .001$, $\eta_p^2 = .176$. Unexpectedly, the interaction between training condition and rating level was not significant, $F(2, 140) = 1.08$, $p = .342$, $\eta_p^2 = .015$, and the three-way interaction was also not significant, $F(2, 140) = 1.02$, $p = .363$, $\eta_p^2 = .014$. See the left panel of Figure 1 for the results.

To break down the interaction effect between measurement time and training condition, we further conducted two separate rmANOVAs with rating level (high vs. low) and training condition (go vs. no-go vs. untrained) as within-subject factors for the pre-training and the post-training ratings, respectively. The rmANOVA on pre-training ratings showed a significant main effect of rating level, $F(1, 70) = 647.3$, $p < .001$, $\eta_p^2 = .902$, indicating that the high-rated items ($M = 54.0$, $SE = 2.12$) were indeed rated as more attractive than the low-rated items ($M = -10.6$, $SE = 2.81$). More important, the main effect of training condition was not significant, $F(1.76, 123.2) = 1.98$, $p = .149$, $\eta_p^2 = .027$. The interaction between rating level and training condition was also not significant, $F(1.79, 125.7) = 0.076$, $p = .910$, $\eta_p^2 = .001$. This result confirmed that

the selection procedure succeeded in creating sets of images that were matched on the average ratings before the training.

For the rmANOVA on post-training ratings, the main effect of rating level remained significant, $F(1, 70) = 303.2, p < .001, \eta_p^2 = .812$. Moreover, the main effect of training condition became significant, $F(1.69, 118.3) = 15.3, p < .001, \eta_p^2 = .179$. The interaction effect was not significant, $F(2, 140) = 1.07, p = .347, \eta_p^2 = .015$. Pairwise comparisons showed that the rating of no-go food items ($M = 11.8, SE = 2.74$) was significantly lower than the rating of go food items ($M = 19.4, SE = 2.43$), $M_{\text{diff}} = -7.63, t(70) = -4.76, p < .001$ (p values were adjusted for multiple comparisons using Bonferroni correction), and untrained items ($M = 17.6, SE = 2.49$), $M_{\text{diff}} = -5.80, t(70) = -3.69, p = .001$, while the difference between go and untrained items did not reach statistical significance, $M_{\text{diff}} = 1.83, t(70) = 1.68, p = .294$. Combined, these results suggested that while before the training, the ratings of all food items were equivalent; no-go foods were rated as less attractive than both untrained and go foods after the training. This decrease in evaluation for no-go foods is the devaluation effect.

Devaluation and Valuation Indices. Although on average, the ratings of food items in three training conditions were matched before the training, this does not mean that they were rated exactly the same for each participant. To account for the small variations in pre-training ratings and get a more accurate estimation of the devaluation effect, for each participant, we calculated a difference score by subtracting the average pre-training rating from the average post-training rating (i.e., difference score = post-training rating - pre-training rating) for each training condition. A negative value thus indicates decrease in rating while a positive value indicates increase. Although the rating level does not seem to moderate the effect, this factor is still of theoretical interest. We therefore calculated the difference scores for high-rated and low-

rated items separately. A devaluation index for no-go items was further calculated by subtracting the difference score of no-go items from that of untrained items (i.e., devaluation index = difference score of untrained items – difference score of no-go items). A higher devaluation index stands for a larger devaluation effect (i.e., more decrease in liking for no-go items in comparison to untrained items). For exploratory reasons, a valuation index for go items was similarly calculated by subtracting the difference score of untrained items from that of go items (i.e., valuation index = difference score of go items – difference score of untrained items). Note that we defined devaluation effect as the evaluation of no-go items being lower than both go items and items that were not used in the training (i.e., untrained items), to rule out alternative explanations (see Chen et al., 2016). For the calculation of the devaluation index, only untrained items were used as the baseline, as the difference between no-go and go items could partially be explained by the go valuation effect. For a summary of the devaluation and valuation indices for high- and low-rated items, see Tables 1 and 2.

Stop-Signal Reaction Time (SSRT). To estimate the SSRT, we used the block-based integration approach recommended by Verbruggen and colleagues (2013). More specifically, SSRT was calculated for each block with the integration approach (for details on the calculation, see Verbruggen et al., 2013), and the average of these four SSRTs was then used as an index of individual inhibition capacity. Two participants had SSRT 3SDs away from the sample mean (one in each direction). Closer inspection of raw data revealed that one participant slowed down significantly in two blocks and had SSRTs below 100 ms in those two blocks, while the other participant did not stop at all in one block and had SSRT above 900 ms in one block. These two participants were therefore excluded. Of the remaining 69 participants, mean go RT ($M = 526.9$ ms, $SD = 60.5$) was significantly longer than the mean failed-stop RT ($M = 453.4$ ms, $SD = 59.6$),

$t(68) = 28.08$, $p < .001$, validating the independence assumption of the race model (Verbruggen

& Logan, 2009). Direction errors ($M = 1.39\%$, $SD = 1.37\%$) and miss rates ($M = 0.92\%$, $SD =$

1.16%) were low. The average stop rate was 49.5%, $SD = 1.85\%$, which is close to the pre-

determined 50%. The average SSD was 259.5 ms, $SD = 82.7$, and the average SSRT was 263.1

ms, $SD = 32.8$.

Correlation Between Devaluation and Inhibition. For the correlation between SSRT

and the devaluation index of high-rated no-go items, 2 out of 69 participants were identified as

bivariate outliers (Cook's distance > 4/N, exclusion criterion not preregistered) and removed,

leading to the predicted significant Pearson's $r(67) = .242$, $p = .048$ (see the left panel of Figure

2). For the correlation between SSRT and the devaluation index of low-rated no-go items, 5 out

of 69 participants were identified as bivariate outliers (Cook's distance > 4/N). This led to non-

significant correlation for low-rated items, Pearson's $r(64) = -.186$, $p = .140$ (see the left panel of

Figure 3).

Discussion

In Experiment 1, we investigated the influence of GNG training on food evaluation and

how this devaluation effect induced by GNG training may be influenced by the attractiveness of

food items and individual inhibition capacity. Overall, we found that no-go items were rated as

less attractive than both go items and items not used in the training. This devaluation effect was

observed for both high-rated and low-rated items. Furthermore, as predicted, a positive

correlation between this devaluation effect and inhibition capacity (i.e., SSRT) was found, while

the correlation between inhibition and devaluation for low-rated items was not statistically

significant. Since the observed effects were small (e.g., correlation between inhibition and

devaluation for high-rated items) and an unexpected devaluation effect for low-rated items was observed, we carried out a replication in Experiment 2.

Experiment 2

Experiment 2 is a replication and extension of Experiment 1. In addition to measuring participants' inhibition capacity, we also measured the other two components of executive control, namely the *updating* and *shifting* components for exploratory reasons. The methods and results concerning the updating and shifting components can be found in the supplementary material. Since we included more measurements of executive control, the experiment was divided into two sessions, with the executive control tasks on session 1 and the GNG training tasks on session 2. Session 2 occurred 1 to 5 day(s) after session 1. We expected to replicate the devaluation effect for high-rated food items, and more important, the positive correlation between inhibition (i.e., SSRT) and the devaluation effect for high-rated no-go items. To encourage conservative analyses, we also preregistered an additional exclusion criterion (i.e., Cook's distance > 4/N) for excluding bivariate outliers. The preregistration can be found at osf.io/cz8zr/.

Method

Sample Size. With a correlation of .242 as the expected effect size (based on Experiment 1), a power analysis indicated 104 participants for achieving 80% power with a one-tailed correlation test when 5% was used as the alpha level. We therefore decided to recruit 110 participants, which exceeded the required sample size and allowed for potential exclusion.

Participants. As in Experiment 1, participants were recruited via the university's online participation system and received course credit or money after the experiment.

Procedure. The current experiment was conducted in two sessions, with the second session scheduled 1 to 5 day(s) after the first session. Since the GNG training is often referred to as inhibitory control training in the literature and sometimes argued to increase the top-down inhibitory control capacity (Houben, & Jansen, 2011), receiving the GNG training first may influence participants' subsequent performance in the stop-signal task (although the evidence available till now seems to argue against this account; see Veling et al., 2017). To reduce such potential influence, we therefore first measured participants' inhibition capacities in Session 1, and then gave them the GNG training in Session 2.

Session One. During the first session, participants received the stop-signal task, which was exactly the same as the one used in Experiment 1. In addition to the stop-signal task, participants also received the number-letter task and the 2-back task to measure their shifting and updating capacity, respectively. The order of these three executive function tasks was counterbalanced across participants. For the procedures and results of the number-letter task and the 2-back task, see the supplementary material. At the end of the session participants filled out the restraint eating scale, and reported demographics such as whether they were currently on a diet, height and weight.

Session Two. During the second session, participants first received unrelated questionnaires for another study, which took around 20 minutes. After a 5-minute break, the current research started. The procedure was exactly the same with Experiment 1 (from *Rating Before the Training* to *Memory Recall*). At the end of the session, participants reported the last time they consumed food, whether they had lunch or not and their current hunger level.

Results

Participants. One hundred and nine participants participated in the study. Three

participants had accuracies on go or no-go trials 3*SD*s below the sample mean and below 90%.

These three participants were excluded (preregistered exclusion criterion). The final sample

consisted of 106 participants (30 males, 76 females, M_{age} = 23.2 years, *SD* = 6.46). On average,

participants finished the second session 2.26 days after the first session (*SD* = 1.62; two

participants participated in the second session 8 and 10 days after the first session, respectively.

Excluding these two participants did not change the results).

Performance in GNG Training. The average accuracy on go trials was *M* = 99.2%, *SD*

= 1.68%; the average no-go accuracy was *M* = 91.9%, *SD* = 7.18%. The average go reaction time

was *M* = 415.3 ms, *SD* = 57.3.

Stimulus Evaluation. A rmANOVA with rating level, measurement time and training

condition as within-subject factors and the average rating as dependent variable was conducted.

As in Experiment 1, all three main effects were significant: rating level, $F(1, 105) = 640.7$, *p*

< .001, $\eta_p^2 = .859$; measurement time, $F(1, 105) = 38.5$, $p < .001$, $\eta_p^2 = .268$; training condition,

$F(2, 210) = 20.4$, $p < .001$, $\eta_p^2 = .162$. Further replicating the results of Experiment 1, the

interaction between rating level and measurement time was significant, $F(1, 105) = 54.5$, *p*

< .001, $\eta_p^2 = .342$; the interaction between measurement time and training condition was also

significant, $F(2, 210) = 22.2$, $p < .001$, $\eta_p^2 = .174$. The interaction between training condition

and rating level was not significant, $F(2, 210) = 0.592$, $p = .554$, $\eta_p^2 = .006$, and the three-way

interaction effect was also not significant, $F(2, 210) = 0.604$, $p = .547$, $\eta_p^2 = .006$. See the right

panel of Figure 1 for the results.

As in Experiment 1, we subsequently conducted two rmANOVAs, on the pre-training

and post-training ratings separately. As would be expected, before the training there was a

significant difference between high- and low-rated items, $F(1, 105) = 750.1$, $p < .001$, $\eta_p^2 = .877$, while the main effect of training condition, $F(1.82, 191.3) = 0.264$, $p = .748$, $\eta_p^2 = .003$, and the interaction effect, $F(1.85, 193.7) = 0.409$, $p = .648$, $\eta_p^2 = .004$, were not significant. The main effect of rating level remained significant after the training, $F(1, 105) = 490.8$, $p < .001$, $\eta_p^2 = .824$, and the main effect of training condition also became significant, $F(2, 210) = 21.6$, $p < .001$, $\eta_p^2 = .171$, while the interaction effect was again not significant, $F(2, 210) = 0.602$, $p = .549$, $\eta_p^2 = .006$. Therefore, the overall pattern on stimulus evaluation replicates the findings of Experiment 1. Pairwise comparisons suggested that no-go items ($M = 3.11$, $SE = 2.46$) were rated as less attractive than both go items ($M = 9.96$, $SE = 2.47$), $M_{diff} = -6.85$, $t(105) = -6.57$, $p < .001$, and untrained items ($M = 6.34$, $SE = 2.59$), $M_{diff} = -3.24$, $t(105) = -2.85$, $p = .005$. Unlike Experiment 1, the go items were rated more positively than untrained items, $M_{diff} = 3.62$, $t(105) = 3.84$, $p < .001$.

Stop-Signal Reaction Time (SSRT). One participant had an invalid value of SSRT (~ 2500 ms). Closer inspection of the data suggested that this participant hardly stopped at all in the first three blocks (p(stop) = 0.0%, 5.0% and 5.0%, respectively) and was therefore excluded. One participant showed SSRT more than 3SDs above the sample mean and was excluded (criterion preregistered). Note that a third participant, with SSRT 2.8 SDs above sample mean, was also excluded. Although his/her SSRT fell in the 3SDs range, inspection of the data showed that he/she made substantial errors on go trials (both direction errors and miss rates were above 20%; compare this to the performance of the remaining participants below), and stopped on only 23.75% of the stop trials. Of the remaining 103 participants, mean go RT was 541.0 ms ($SD = 72.7$), which was significantly longer than the failed-stop RT ($M = 464.1$ ms, $SD = 62.3$), $t(102) = 28.9$, $p < .001$. Direction errors ($M = 1.78\%$, $SD = 2.09\%$) and miss rates ($M = 1.38\%$, $SD = 1.90\%$)

were again low. On average, participants stopped on 49.5% ($SD = 2.71\%$) of the stop trials. The

average SSD was 262.3 ms ($SD = 80.1$), and the average SSRT was 276.9 ms ($SD = 35.1$).

Correlation Between Inhibition and Devaluation. Before conducting the main

correlation analysis, we identified outliers step by step as specified in our preregistration.

Univariate outliers that are 3SDs from the sample mean were first removed. This led to the

exclusion of 3 more participants since their devaluation index for high-rated no-go items was

3SDs away from sample mean (2 above and 1 below). Bivariate outliers (Cook's distance > 4/N,

1 identified) were then removed. Ninety-nine participants remained in the analysis. Contrary to

our prediction and the finding in Experiment 1, the correlation was not significant, $r(99) = -.018$,

$p = .861$ (see the middle panel of Figure 2). The same procedure was used for calculating the

correlation between inhibition and devaluation for low-rated no-go items. Unexpectedly and

contrary to the finding in Experiment 1, the correlation reached statistical significance, $r(95)$

$= .220$, $p = .033$, so that participants with lower inhibition capacity (i.e., longer SSRT) showed

stronger devaluation effect for low-rated items (see the middle panel of Figure 3).

Discussion

In Experiment 2, we predicted and replicated the devaluation effect for high-rated items.

Although unexpected, the devaluation effect for low-rated items was again found, replicating the

result of Experiment 1. Furthermore, a valuation effect was found for go items, where such a

valuation effect failed to reach statistical significance in Experiment 1.

Experiment 2 failed to replicate the positive correlation between inhibition and

devaluation for high-rated items as observed in Experiment 1, and quite unexpectedly (and

inconsistent with the finding in Experiment 1), a positive correlation between inhibition and

devaluation emerged for low-rated items. Given these inconsistencies in the results of

Experiments 1 and 2, and the similarities in experimental procedures, in the next part we combined data from two experiments and carried out combined and exploratory analyses.

Combined and Exploratory Analyses

Stimulus Evaluation. To evaluate whether the strength of no-go devaluation and go valuation effects differs between two experiments and between high-rated and low-rated items, we combined the data from two experiments. First, the devaluation indices were submitted to a rmANOVA with rating level (high vs. low) as the within-subject factor and the experiment as the between-subject factor. None of the effects reached statistical significance, Fs < 1.38, ps > .241. Since the conventional NHST approach has difficulty in interpreting non-significant results, a Bayesian rmANOVA was carried out with JASP (JASP Team, 2016). For all effects, the Bayes factor (i.e., BF) was below 1/3, which provided moderate support for the null model (Dienes, 2014; for the output of the Bayesian rmANOVA, see the supplementary material), suggesting that the devaluation effect was similar in both experiments and also did not differ for high-rated and low-rated items.

Since no difference in the strength of devaluation effects was found between the high-rated and low-rated items and between Experiment 1 and 2, we collapsed the four devaluation indices. The overall effect size of the devaluation effect is $d = 0.346$, 95% CI [0.194, 0.497]. A same rmANOVA was conducted on the go valuation indices, and none of the effects reached statistical significance, Fs < 1.88, ps > .172. Bayesian rmANOVA on valuation indices gives BFs below 1/3, again suggesting that the go valuation effects did not differ between high- and low-rated items or between the two experiments. The overall effect size of the go valuation effect is $d = 0.312$, 95% CI [0.161, 0.463].

Contrary to our expectations, the low-rated items were devalued to the same extent as the high-rated items. This raised the question of to what extent the high-rated items were indeed more impulse-invoking that the low-rated items. To explore this, we divided the GNG training into two parts, trials on which high-rated items were presented, and trials on which low-rated items were presented. Performance in these two parts was then compared. As can be seen from Table 3, in both Experiments 1 and 2, participants responded to high-rated items more quickly than to low-rated items, and their accuracy in responding to high-rated items was also higher. Furthermore, they made more mistakes when they needed to inhibit their responses to high-rated items in comparison to low-rated items, although this difference on no-go accuracies did not reach statistical significance in Experiment 1. Overall, these results suggest that in line with our expectations, high-rated items indeed invoked more impulses (in this case, to respond to the food and difficulty to stop toward food) than low-rated items did.

Since the food-pics database (Blechert et al., 2014) also provided caloric density information for all the food images, we also extracted this information for the selected food images. For each participant, the average caloric density was calculated for the high-rated and low-rated items separately. Paired-sample t tests suggested that in both experiments, the average caloric density was higher for the low-rated items than the high-rated items (Experiment 1: low-rated items, $M = 290.7$, $SD = 31.7$; high-rated items, $M = 267.4$, $SD = 28.3$, $t(70) = 3.58$, $p < .001$, $d_z = 0.425$, 95% CI [0.181, 0.667]. Experiment 2: low-rated items, $M = 290.4$, $SD = 34.3$; high-rated items, $M = 261.1$, $SD = 35.7$, $t(105) = 4.77$, $p < .001$, $d_z = 0.463$, 95% CI [0.262, 0.662]). This finding is in line with the overall pattern observed in the food-pics database. That is, the low calorie-dense foods in the food-pics database received higher ratings of palatability and desire to eat than high calorie-dense foods (Blechert et al., 2014). This result may be due to the inclusion

of certain food items (e.g., low calorie-dense food items that are highly appetitive) in the database and in the current study.

In the current study, high- and low-rated items were used in the same training, whereas in previous work, rating level was often manipulated between subjects (e.g., Chen et al., 2016; but see Veling et al., 2008; Experiment 2). Previous work employing a similar general procedure and same stimuli has found that when the training contained only high-rated items, devaluation effect was observed (Experiment 4a; Chen et al., 2016), whereas when the training contained only low-rated items (Experiment 6; Chen et al., 2016) this effect was absent. To explore whether these different results could be explained by differences in learning the stimulus-response contingencies, we carried out signal detection analysis on the memory recall data from the current study (i.e., within-subject manipulation of rating level) and previous work (Experiments 4a and 6, Chen et al., 2016; between-subjects manipulation of rating level). The signal detection analysis is superior to analyses on overall accuracies, as it disentangles two distinct components, namely discriminability and response bias (Macmillan, & Creelman, 2004). In this case, the discriminability parameter shows to what extent participants could discriminate the go items from the no-go items, and is thus an indicator of learning. The response bias parameter, on the other hand, indicates participant's general tendency to say whether an item was a go or no-go item in the memory recall task. For the data analysis steps and the complete result, see the supplementary material (Table S4). Overall, we found that when rating level was manipulated between subjects, participants showed higher discriminability towards high-rated items than low-rated items (effect size Hedge's g ranged from 0.483 to 0.535), whereas the response bias parameter was comparable (Hedge's g ranged from -0.182 to -0.016). The pattern is opposite when rating level was manipulated within subjects (as in the current study):

participants showed slightly higher discriminability towards high-rated items than low-rated

items (Hedge's g ranged from 0.081 to 0.110), and much stronger negative response bias towards

high-rated items than low-rated items (Hedge's g ranged from -0.790 to -0.479). Overall, these

results suggest that participants learnt the contingencies better for high-rated items in between-

subjects design, but this advantage in learning for high-rated items is much less pronounced in

within-subject design. On the contrary, when high-rated and low-rated items were included in

one training, participants had a strong response bias to say a high-rated item was a go item,

whereas this bias is much smaller in a between-subjects design. The difference in learning as a

function of the procedural difference may be explained by the amount of attention participants

devote to the training. Recent work on stimulus-stop learning showed that participants learned

the associations between stimulus and stopping only when they paid sufficient attention to the

task (Best, Lawrence, Logan, McLaren, & Verbruggen, 2016). By presenting both appetitive and

less appetitive stimuli in one training, participants may have paid more attention to the whole

training and therefore have learnt the contingencies for low-rated items as well. When rating

level was manipulated between subjects, this may lead to low attention to the training when

neutral or low-rated stimuli were presented (i.e., because these stimuli are relatively

uninteresting), and the devaluation effect was thus absent.

Correlation Between Inhibition and Devaluation. Before combining the data, it is useful to

note the procedural difference between Experiments 1 and 2. That is, while in Experiment 1

participants finished both the stop-signal task and the go/no-go training in one session, in

Experiment 2 these two tasks were separated into two sessions, with on average 2.26 days in

between. In hindsight, one potential explanation for the diverging findings is that one's inhibition

capacity may fluctuate, and inhibition capacity measured temporally close to the training may be

more predictive of the training's effectiveness. Although plausible, we think this procedural difference could not easily explain the difference between Experiments 1 and 2, since inhibition capacity is often conceptualized as an individual trait (e.g., the stop-signal reaction time has been shown to correlate with self-report measurement of impulsivity, Logan, Schachar, & Tannock, 1997) and has genetic origins (Friedman, Miyake, Young, DeFries, Corley, & Hewitt, 2008), and therefore is assumed to be relatively stable over time (test-retest reliability of the stop-signal reaction time is .65 on two session 8.6 days apart from each other; Weafer, Baggott, & de Wit, 2013).

We then combined data from two experiments. For the correlation between SSRT and the devaluation effect for high-rated items, 166 participants remained in the combined analysis (see the right panel of Figure 2). The correlation was not statistically significant, $r(166) = .059$, $p = .446$. Bayesian correlation analysis gave a BF of 0.129 (Beta* prior width = 1; JASP team, 2016), which provided support for the null hypothesis. For the correlation between SSRT and the devaluation effect for low-rated items, 159 participants remained in the combined analysis (see the right panel of Figure 3). The correlation was also not significant, $r(159) = .033$, $p = .682$, BF = 0.108, which also supported the null hypothesis. Overall, the combined analysis on Experiments 1 and 2 suggested that there was no correlation between individual inhibition capacity and the strength of devaluation effect.

General Discussion

In two experiments, we investigated the effect of GNG training on food evaluation and tested whether this effect is influenced by the attractiveness of food items and individual inhibition capacity. In two experiments, we consistently found a devaluation effect for no-go items, for both high- and low-rated foods, and a valuation effect for go items (the effect was not

significant in Experiment 1 though), again for both high- and low-rated foods. The predicted

correlation between inhibition and devaluation for high-rated foods was found in Experiment 1,

but not in Experiment 2 and not in a combined analysis of Experiments 1 and 2. The correlation

between inhibition and devaluation for low-rated foods was also not significant in the combined

analysis. We discuss the implications of these results below.

Stimulus Evaluation. As expected, and consistent across two experiments, we found that

high-rated no-go items were evaluated as less attractive in comparison to both go and untrained

items. However, contrary to our expectations, but consistent across two experiments, the

devaluation effect was also found for low-rated items. Furthermore, the strength of the

devaluation effect did not differ between high- and low-rated food items.

This devaluation effect for low-rated items seems not in line with previous findings that

showed GNG training is more effective in reducing the evaluations of positive or appetitive

stimuli (Veling et al., 2008; Chen et al., 2016), or in samples for which the food stimuli were

more appetitive (Houben, 2011; Houben & Jansen, 2011; Lawrence et al., 2015b; Veling et al.,

2011; Veling et al., 2013a). Our exploratory analysis showed that participants responded faster to

high-rated items than to low-rated items in GNG training, suggesting that high-rated items were

indeed more impulse invoking. However, the low-rated foods had higher caloric density than the

high-rated foods (in line with the general pattern observed in the food-pics database; Blechert et

al., 2014), suggesting that even the low-rated items might still be appetitive enough to elicit

impulsive responses in GNG training, which were inhibited upon the presentation of no-go cues.

Furthermore, our exploratory signal detection analysis showed that participants learnt the

stimulus-response contingencies to similar degrees for high-rated and low-rated items,

suggesting that presenting low-rated items together with high-rated items may have increased

attention towards these items and facilitated learning. To tackle this problem and further explore

the role of attention in such training, future research may directly manipulate attention and

observe its influence on learning and behavioral outcomes.

In addition to the no-go devaluation effect, we also found a go valuation effect. That is,

the evaluation of go items was higher than that of untrained items, and the strength of this effect

did not differ between rating levels. This go valuation effect was consistent across two

experiments (although it did not reach statistical significance in Experiment 1), and was also in

line with our previous findings in which a similar training procedure was used (Chen et al., 2016,

Experiments 4a and 6). This result may be explained by the recent work on the interaction effect

of action and valence (Guitart-Masip, Duzel, Dolan, & Dayan, 2014). More specifically, this line

of work showed that reward and behavioral activation (e.g., go responses), and punishment and

behavioral inhibition (i.e., no-go responses) are closely coupled and have some overlapping

neurological and physiological underpinnings (Guitart-Masip et al., 2014). Since the anticipation

of go actions activates similar brain areas as the response to reward, responding frequently may

engage the reward region on the go trials, thereby increasing the evaluation of go foods. Future

work can employ neuroimaging tools such as fMRI to elucidate the nature and the underlying

mechanism of such a go valuation effect.

Correlation Between Inhibition and Devaluation. In Experiment 1, we observed the

predicted association between inhibition and devaluation for high-rated foods, and a non-

significant correlation for low-rated foods. However, an opposite pattern was found in

Experiment 2: a non-significant correlation for high-rated foods, and a positive correlation (albeit

also small) for low-rated foods. When combining the data from both experiments, none of the

correlations was statistically significant, and Bayesian correlation analyses further showed moderate support for the null hypothesis.

One potential explanation for the absence of correlation can be that different types of inhibition are measured with the stop-signal task and are engaged on the no-go trials in GNG training. In the stop-signal task, participants are required to indicate the direction of an arrow unless a stop-signal is played; a *general inhibition capacity* is assumed to be engaged in order to successfully stop. In GNG training, however, the tendency to respond is triggered by the food item presented, and in order to inhibit this impulse triggered by foods, a *food-specific inhibition capacity* may be required. Houben and colleagues (2014) measured participants' general inhibition capacity with a stop-signal task with nonfood items (like the one used in the current research), and food-specific inhibition capacity with a modified stop-signal task in which food items were presented, and found that food-specific inhibition capacity, but not the general inhibition capacity, was correlated with body weight. Note that although previous studies have shown that body weight is related to the general inhibition capacity, most such studies relied on the comparison between overweight or obese individuals and healthy controls (Lavagnino et al., 2016). In a sample with normal body weight, as the sample in the current study, the general inhibition capacity may not be strongly related to the food-specific inhibition capacity and may not be very predictive of the effectiveness of GNG training.

One may also question the assumption that individuals with lower (general or food-specific) inhibition capacity experience more conflict on the no-go trials in GNG training, and hence benefit more from the training. Like previous studies on GNG training (e.g., Chen et al., 2016), participants performed quite well in GNG training in the current study, and made very few errors. The training is not very challenging: participants were instructed to press a key or not

based on the tones they heard, and the relatively long response window (i.e., 900ms) provided

ample time to respond. It is therefore possible that this task did not pose a strong challenge on

the inhibition capacities of our participants. Hence, the strength of the conflict experienced by

the participants may be comparable regardless of the individual inhibition capacities, since it is

well within all participants' capabilities to withhold their responses in such an easy training task.

      Alternative Explanations and Future Directions. The two predictions of the BSI

theory rest on the core assumption that conflict between an approach tendency toward a stimulus

and response inhibition leads to stimulus devaluation. Although previous work indeed suggests

that conflicts are registered as aversive signals (Botvinick, 2007; Dreisbach, & Fischer, 2015), it

remains unclear whether the devaluation effect observed in GNG training is caused by such

aversive conflict signals. For instance, Wessel and colleagues showed that presenting items on

incongruent trials in a Simon paradigm did not lead to devaluation (Wessel et al., 2014,

Experiment 3), and concluded that devaluation might be caused by action stopping, but not

response conflict. Successful action stopping can have a global motor suppressive effect (Wessel,

& Aron, 2017) and this inhibitory effect may spill over to the amygdala (Berkman, Burklund, &

Lieberman, 2009), thereby reducing the affective responses towards the stimulus associated with

action stopping. In addition to action stopping, other findings suggest that inhibition in general

may have negative affective consequences for the inhibited stimulus, as inhibition of attention

towards a stimulus (Fenske, & Raymond, 2006) and inhibition of stimulus memory (De Vito, &

Fenske, 2017) have both been shown to devalue the stimulus. These studies mainly used neutral

stimuli (e.g., abstract patterns, unfamiliar faces, neutral words and objects) and the devaluation

effect has also been consistently observed. To explain such effects, inhibition (attention or

memory) has also been directly linked to devaluation without recourse to response conflict. Such

a pure devaluation-by-inhibition account (i.e., an account that does not assume a role for response conflict) would predict devaluation effect for positive and neutral stimuli alike, and perhaps also similar effects for individuals with different inhibition capacities.

Related, the devaluation-by-inhibition account also provides a different prediction than BSI theory (Veling et al., 2008) on the generalizability of the devaluation effect to stimuli other than food. While effects of GNG training have been found on evaluations of alcoholic drinks (Houben, Nederkoorn, Wiers, & Jansen, 2011; Houben, Havermans, Nederkoorn, & Jansen, 2012), faces (Kiss, Raymond, Westoby, Nobre, & Eimer, 2008; Doallo, Raymond, Shapiro, Kiss, Eimer, & Nobre, 2012) and sexual pictures (Ferrey, Frischen, & Fenske, 2012), most of the stimuli used were relatively rewarding. According to BSI theory, one prerequisite of the devaluation effect is for the stimulus to evoke an approach tendency; hence the effect may generalize to other types of stimuli as long as they are impulse evoking. The devaluation-by-inhibition account, on the other hand, does not make such an assumption and may predict generalization of the devaluation effect to neutral or even negative stimuli (see e.g., Frischen et al., 2012; but see Chen et al., 2016). Till now, many of the studies using GNG training have examined effects on positive stimuli (e.g., palatable food, alcoholic drinks etc.), given the apparent applied value of devaluing such positive stimuli. Future work may include more neutral or even negative stimuli in GNG training to gain more theoretical understanding of the underlying mechanisms.

In addition to investigating devaluation of neutral stimuli, neuroimaging measures may also be helpful to test these different accounts. For instance, it remains interesting to examine how the brain regions related to motor excitation, reward processing, and response inhibition may interact (for a map of brain regions potentially associated with the underlying mechanisms

of such trainings, see Stice, Lawrence, Kemps, & Veling, 2016). Future work may also directly

measure the strength of conflict in GNG training, for instance conflict-related activity in the

anterior cingulate cortex (Kerns, Cohen, MacDonald, Cho, Stenger, & Carter, 2004) and test its

role in devaluation. Note that these different accounts are not necessarily mutually exclusive, and

different processes may be engaged at different stages of the training (Veling et al., 2017). A

neuroimaging approach may hence also be useful in providing more insight into the potential

changes in the underlying mechanisms as the training progresses.

Furthermore, future research may also consider that inhibition process might influence

the consumption of high-calorie food either directly (i.e., inhibit the consumptive behavior itself)

or indirectly (i.e., reduce the craving responses that underlie consumption; see Hall, 2016). By

using explicit food evaluation as the dependent measurement, we mainly focused on the indirect

route. Note that Houben (2011) used the amount of consumption as the dependent measurement

and showed that inhibition training was more effective for individuals with lower (general)

inhibition capacity. In light of this difference, future work may take into account the differences

between various measurements and explore how these different measures may be selectively

influenced by the training and individual inhibition capacity. For instance, to measure food

evaluations, future work may use either explicit self-report (as in the current study) or more

indirect measurements (e.g., implicit association test and affective priming task, De Houwer,

Teige-Mocigemba, Spruyt, & Moors, 2009), as the predictive validities of explicit and implicit

measures may differ depending on the context and situation (Friese, Hofmann, & Wänke, 2008).

From an applied perspective, it is important to use behavioral measures in addition to food

evaluation (explicit or implicit), such as the amount of consumption, food choice etc., to better

understand the multiple facets of real eating behaviors, and the influences of GNG training on these different aspects.

Conclusion

In conclusion, across two experiments we predicted and consistently found that inhibiting responses to appetitive foods in GNG training leads to lower evaluation of these foods. An unexpected, but also consistent devaluation effect was found for less appetitive foods as well. Exploratory signal detection analysis on memory recall data suggests that participants may have learnt the contingencies between stimulus and response equally well for appetitive and less appetitive food items in the current setting. Regarding the question raised in the title of this article, do impulsive individuals benefit more from food inhibition training, the answer is no, individuals with lower (general) inhibition capacity do not show a larger devaluation effect. These two findings are thus not in line with the predictions made by BSI theory. Future work may measure food-specific inhibition capacity and investigate whether it better predicts the effectiveness of training, or measure the strength of conflict directly to test the assumption of BSI theory that devaluation in GNG training is driven by response conflict between impulses and response inhibition. In sum, the current research reinforces the notion that inhibiting responses toward appetitive foods in GNG training may be a promising way to help people eat more healthily in this obesogenic world.

References

Adams, R. C., Lawrence, N. S., Verbruggen, F., & Chambers, C. D. (2017). Training response

    inhibition to reduce food consumption: Mechanisms, stimulus specificity and appropriate

    training protocols. *Appetite, 109*, 11-23.

Allom, V., Mullan, B., & Hagger, M. (2015). Does inhibitory control training improve health

    behaviour? A meta-analysis. *Health Psychology Review*, 1-38.

Berkman, E. T., Burklund, L., & Lieberman, M. D. (2009). Inhibitory spillover: Intentional

    motor inhibition produces incidental limbic inhibition via right inferior frontal cortex.

    Neuroimage, 47, 705-712.

Best, M., Lawrence, N. S., Logan, G. D., McLaren, I. P., & Verbruggen, F. (2016). Should I stop

    or should I go? The role of associations and expectancies. *Journal of Experimental*

    Psychology: Human Perception and Performance, 42, 115–137.

Blechert, J., Meule, A., Busch, N. A., & Ohla, K. (2015). Food-pics: an image database for

    experimental research on eating and appetite. *Frontiers in Psychology*, 40.

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives

    on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience, 7*, 356-

    366.

Breslin, P. A. (2013). An evolutionary perspective on food and human taste. *Current Biology, 23*,

    R409-R418.

Brunstrom, J. M., Yates, H. M., & Witcomb, G. L. (2004). Dietary restraint and heightened

    reactivity to food. *Physiology & Behavior, 81*, 85-90.

Chen, Z., Veling, H., Dijksterhuis, A., & Holland, R.W. (2016). How does not responding to appetitive stimuli cause devaluation: Evaluative conditioning or response inhibition? *Journal of Experimental Psychology: General*, *145*, 1687-1701.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135*, 347.

De Vito, D., & Fenske, M. J. (2017). Suppressing memories of words and familiar objects results in their affective devaluation: Evidence from Think/No-think tasks. *Cognition, 162*, 1-11.

Doallo, S., Raymond, J. E., Shapiro, K. L., Kiss, M., Eimer, M., & Nobre, A. C. (2012). Response inhibition results in the emotional devaluation of faces: neural correlates as revealed by fMRI. *Social Cognitive and Affective Neuroscience, 7*, 649-659.

Dreisbach, G., & Fischer, R. (2015). Conflicts as aversive signals for control adaptation. *Current Directions in Psychological Science, 24*, 255-260.

Fedoroff, I. C., Polivy, J., & Herman, C. P. (1997). The effect of pre-exposure to food cues on the eating behaviour of restrained and unrestrained eaters. *Appetite, 28*, 33–47.

Fenske, M. J., & Raymond, J. E. (2006). Affective influences of selective attention. *Current Directions in Psychological Science, 15*, 312-316.

Ferrey, A. E., Frischen, A., & Fenske, M. J. (2012). Hot or not: response inhibition reduces the hedonic value and motivational incentive of sexual stimuli. *Frontiers in Psychology, 3*.

Folkvord, F., Veling, H., & Hoeken, H. (2016). Targeting implicit approach reactions to snack food in children: Effects on intake. *Health Psychology, 35*, 919-922.

Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour. *British Journal of Social Psychology, 47*, 397-419.

Frischen, A., Ferrey, A. E., Burt, D. H., Pistchik, M., & Fenske, M. J. (2012). The affective

consequences of cognitive inhibition: devaluation or neutralization? *Journal of*

Experimental Psychology: Human Perception and Performance, 38, 169-179.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008).

Individual differences in executive functions are almost entirely genetic in origin. *Journal*

of Experimental Psychology: General, 137, 201-225.

Guerrieri, R., Nederkoorn, C., Stankiewicz, K., Alberts, H., Geschwind, N., Martijn, C., &

Jansen, A. (2007). The influence of trait and induced state impulsivity on food intake in

normal-weight healthy women. *Appetite, 49*, 66-73.

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision

making. *Trends in Cognitive Sciences, 18*, 194-202.

Gupta, N., & Aron, A. R. (2011). Urges for food and money spill over into motor system

excitability before action is taken. *European Journal of Neuroscience, 33*, 183-188.

Hall, P. A. (2016). Executive-control processes in high-calorie food consumption. *Current*

Directions in Psychological Science, 25, 91-98.

Harvey, K., Kemps, E., & Tiggemann, M. (2005). The nature of imagery processes underlying

food cravings. *British Journal of Health Psychology, 10*, 49-56.

Herman, C. P., & Polivy, J. (1980). Restrained eating. In A. J. Stunkard (Ed.), *Obesity* (pp. 208-

225). Philadelphia: Saunders

Hofmann, W., Friese, M., & Roefs, A. (2009). Three ways to resist temptation: The independent

contributions of executive attention, inhibitory control, and affect regulation to the impulse

control of eating behavior. *Journal of Experimental Social Psychology, 45*, 431-435.

Houben, K. (2011). Overcoming the urge to splurge: Influencing eating behavior by

    manipulating inhibitory control. *Journal of Behavior Therapy and Experimental Psychiatry,*

    42, 384-388.

Houben, K., Havermans, R. C., Nederkoorn, C., & Jansen, A. (2012). Beer à No-Go: Learning to

    stop responding to alcohol cues reduces alcohol intake via reduced affective associations

    rather than increased response inhibition. *Addiction, 107*, 1280-1287.

Houben, K., & Jansen, A. (2011). Training inhibitory control. A recipe for resisting sweet

    temptations. *Appetite, 56*, 345-349.

Houben, K., & Jansen, A. (2015). Chocolate equals stop: chocolate-specific inhibition training

    reduces chocolate intake and go associations with chocolate. *Appetite, 87*, 318-323.

Houben, K., Nederkoorn, C., & Jansen, A. (2014). Eating on impulse: The relation between

    overweight and food-specific inhibitory control. *Obesity, 22*, E6-E8.

Houben, K., Nederkoorn, C., Wiers, R. W., & Jansen, A. (2011). Resisting temptation:

    decreasing alcohol-related affect and drinking behavior by training response inhibition.

    Drug and Alcohol Dependence, 116, 132-136.

JASP Team (2016). JASP (Version 0.7.5.5)[Computer software].

Kiss, M., Raymond, J. E., Westoby, N., Nobre, A. C., & Eimer, M. (2008). Response inhibition

    is linked to emotional devaluation: Behavioural and electrophysiological evidence.

    Frontiers in Human Neuroscience, 2.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004).

    Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*, 1023-1026.

Kulendran, M., Vlaev, I., Sugden, C., King, D., Ashrafian, H., Gately, P., & Darzi, A. (2014). Neuropsychological assessment as a predictor of weight loss in obese adolescents. *International Journal of Obesity, 38*, 507-512.

Lavagnino, L., Arnone, D., Cao, B., Soares, J. C., & Selvaraj, S. (2016). Inhibitory control in obesity and binge eating disorder: A systematic review and meta-analysis of neurocognitive and neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 68*, 714-726.

Lawrence, N. S., O'Sullivan, J., Parslow, D., Javaid, M., Adams, R. C., Chambers, C. D., Kos, K., & Verbruggen, F. (2015a). Training response inhibition to food is associated with weight loss and reduced energy intake. *Appetite, 95*, 17-28.

Lawrence, N. S., Verbruggen, F., Morrison, S., Adams, R. C., & Chambers, C. D. (2015b). Stopping to food can reduce intake. Effects of stimulus-specificity and individual differences in dietary restraint. *Appetite, 85*, 91-103.

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: a model and a method. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 276-291.

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and inhibitory control. *Psychological Science, 8*, 60-64.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology Press.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Current Directions in Psychological Science, 21*, 8-14.

Nederkoorn, C., Braet, C., Van Eijs, Y., Tanghe, A., & Jansen, A. (2006a). Why obese children

cannot resist food: the role of impulsivity. *Eating Behaviors, 7*, 315-322.

Nederkoorn, C., Smulders, F. T., Havermans, R. C., Roefs, A., & Jansen, A. (2006b).

Impulsivity in obese women. *Appetite, 47*, 253-256.

Nederkoorn, C., Houben, K., Hofmann, W., Roefs, A., & Jansen, A. (2010). Control yourself or

just eat what you like? Weight gain over a year is predicted by an interactive effect of

response inhibition and implicit preference for snack foods. *Health Psychology, 29*, 389 -

393.

Nederkoorn, C., Van Eijs, Y., & Jansen, A. (2004). Restrained eaters act on impulse. *Personality

and Individual differences, 37*, 1651-1658.

Nijs, I. M., Muris, P., Euser, A. S., & Franken, I. H. (2010). Differences in attention to food and

food intake between overweight/obese and normal-weight females under conditions of

hunger and satiety. *Appetite*, *54*, 243-254.

Papies, E., Stroebe, W., & Aarts, H. (2007). Pleasure in the mind: Restrained eating and

spontaneous hedonic thoughts about food. *Journal of Experimental Social Psychology, 43*,

810-817.

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience

Methods, 162*, 8-13.

Phan, U. T., & Chambers, E. (2016). Motivations for choosing various food groups based on

individual foods. *Appetite, 105*, 204-211.

Seibt, B., Häfner, M., & Deutsch, R. (2007). Prepared to eat: How immediate affective and

motivational responses to food cues are influenced by food deprivation. *European Journal

of Social Psychology, 37*, 359-379.

Stice, E., Lawrence, N. S., Kemps, E., & Veling, H. (2016). Training motor responses to food: A

    novel treatment for obesity targeting implicit processes. *Clinical Psychology Review, 49*,

    16-27.

Stirling, L. J., & Yeomans, M. R. (2004). Effect of exposure to a forbidden food on eating in

    restrained and unrestrained women. *International Journal of Eating Disorders, 35*, 59-68.

Veenstra, E. M., & de Jong, P. J. (2010). Restrained eaters show enhanced automatic approach

    tendencies towards food. *Appetite, 55*, 30-36.

Veling, H., Aarts, H., & Papies, E. K. (2011). Using stop signals to inhibit chronic dieters'

    responses toward palatable foods. *Behaviour Research and Therapy, 49*, 771-780.

Veling, H., Aarts, H., & Stroebe, W. (2013a). Stop signals decrease choices for palatable foods

    through decreased food evaluation. *Frontiers in Psychology, 4*, 875.

Veling, H., Aarts, H., & Stroebe, W. (2013b). Using stop signals to reduce impulsive choices for

    palatable unhealthy foods. *British Journal of Health Psychology, 18*, 354-368.

Veling, H., Holland, R. W., & van Knippenberg, A. (2008). When approach motivation and

    behavioral inhibition collide: Behavior regulation through stimulus devaluation. *Journal of

    Experimental Social Psychology, 44*, 1013-1019.

Veling, H., Lawrence, N. S., Chen, Z., van Koningsbruggen, G. M., & Holland, R. W. (2017).

    What is trained during food go/no-go training? A review focusing on mechanisms and a

    research agenda. *Current Addiction Reports, 4*, 35-41.

Veling, H., van Koningsbruggen, G. M., Aarts, H., & Stroebe, W. (2014). Targeting impulsive

    processes of eating behavior via the internet. Effects on body weight. *Appetite, 78*, 102-109.

Verbruggen, F., Chambers, C. D., & Logan, G. D. (2013). Fictitious inhibitory differences how

skewness and slowing distort the estimation of stopping latencies. *Psychological Science,*

24, 352-362.

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends*

in Cognitive Sciences, 12, 418-424.

Verbruggen, F., & Logan, G. D. (2009). Models of response inhibition in the stop-signal and

stop-change paradigms. *Neuroscience & Biobehavioral Reviews, 33*, 647-661.

Wang, G. J., Volkow, N. D., Telang, F., Jayne, M., Ma, J., Rao, M., ... & Fowler, J. S. (2004).

Exposure to appetitive food stimuli markedly activates the human brain. *Neuroimage, 21*,

1790-1797.

Weafer, J., Baggott, M. J., & de Wit, H. (2013). Test–retest reliability of behavioral measures of

impulsive choice, impulsive action, and inattention. *Experimental and Clinical*

Psychopharmacology, 21, 475-481.

Wessel, J. R., & Aron, A. R. (2017). On the Globality of Motor Suppression: Unexpected Events

and Their Influence on Behavior and Cognition. *Neuron, 93*, 259-280.

Wessel, J. R., O'Doherty, J. P., Berkebile, M. M., Linderman, D., & Aron, A. R. (2014).

Stimulus devaluation induced by stopping action. *Journal of Experimental Psychology:*

General, 143, 2316–2329.

World Health Organization. (2016). Obesity and overweight. Retrieved June 15, 2016, from

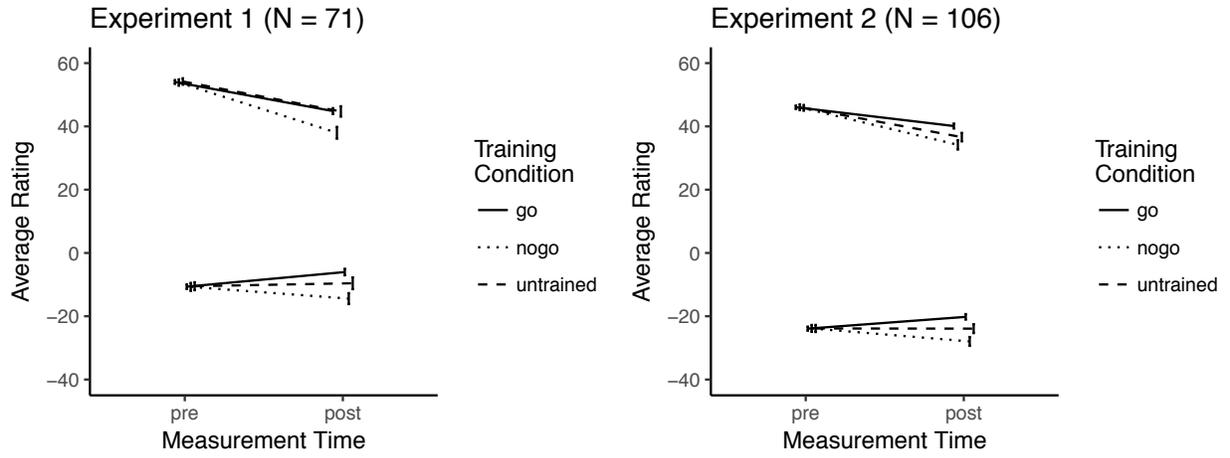http://www.who.int/mediacentre/factsheets/fs311/en/

Figure 1. Evaluations of food items in Experiment 1 (left) and Experiment 2 (right). Error bars stand for within-subject standard errors.

Figure 2. Correlations Between Inhibition Capacity (SSRT) and No-Go Devaluation Effect of High-Rated Items in Experiment 1 (left panel), Experiment 2 (middle panel) and combined (right panel).
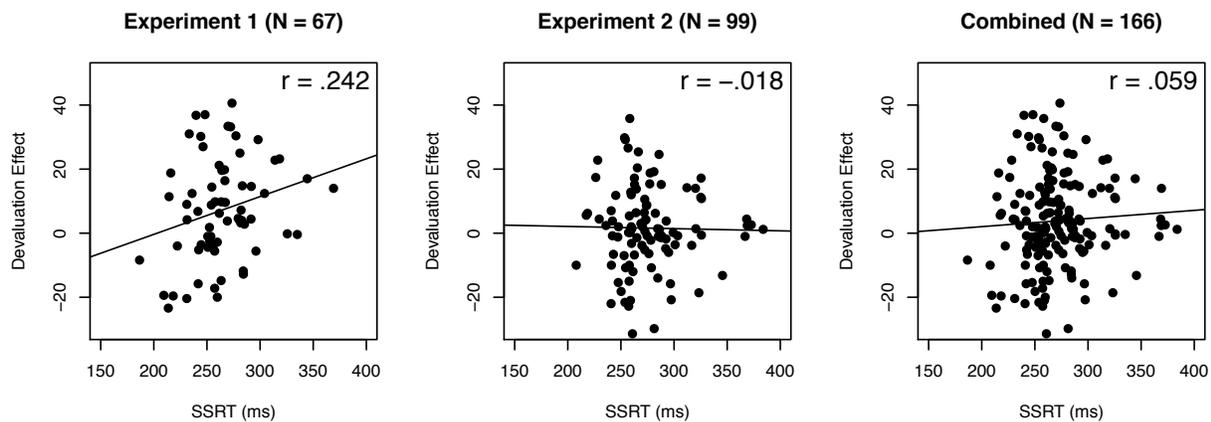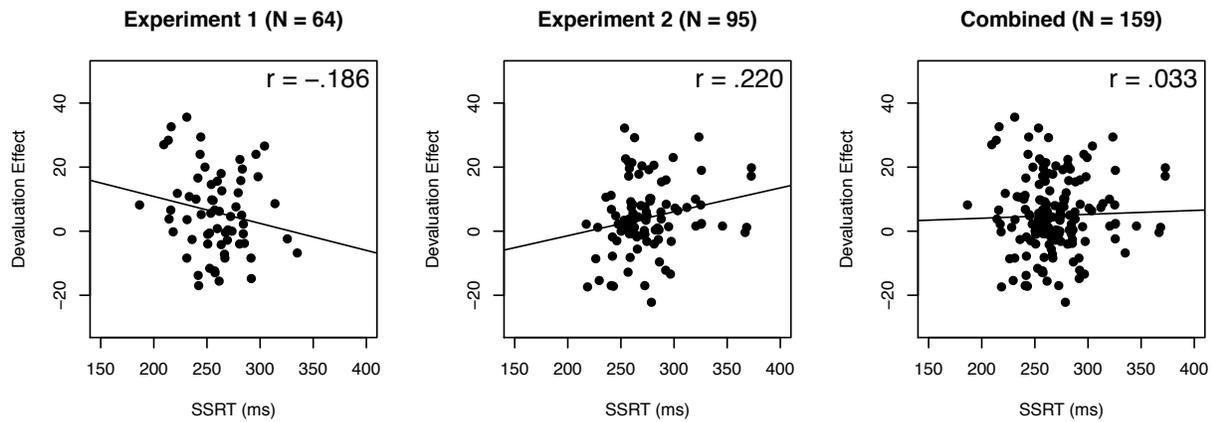
Figure 3. Correlations Between Inhibition Capacity (SSRT) and No-Go Devaluation Effect of Low-Rated Items in Experiment 1 (left panel), Experiment 2 (middle panel) and combined (right panel).

Study	Rating Level	Cohen's d [95% CI]	
Study 1	High	0.368 [0.127, 0.608]	
Study 1	Low	0.272 [0.034, 0.508]	
Study 2	High	0.163 [−0.029, 0.354]	
Study 2	Low	0.281 [0.086, 0.474]	
Combined		0.346 [0.194, 0.497]	
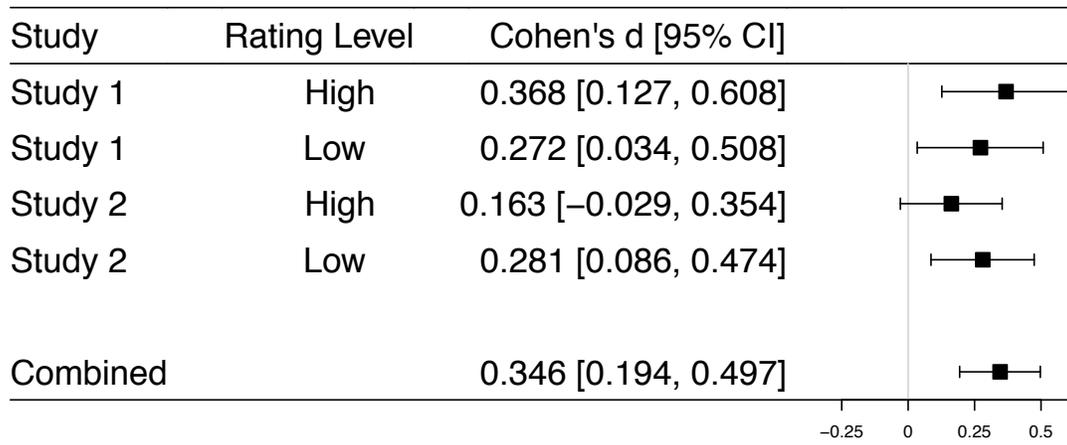
Table 1. Effect Sizes of No-Go Devaluation Effects.

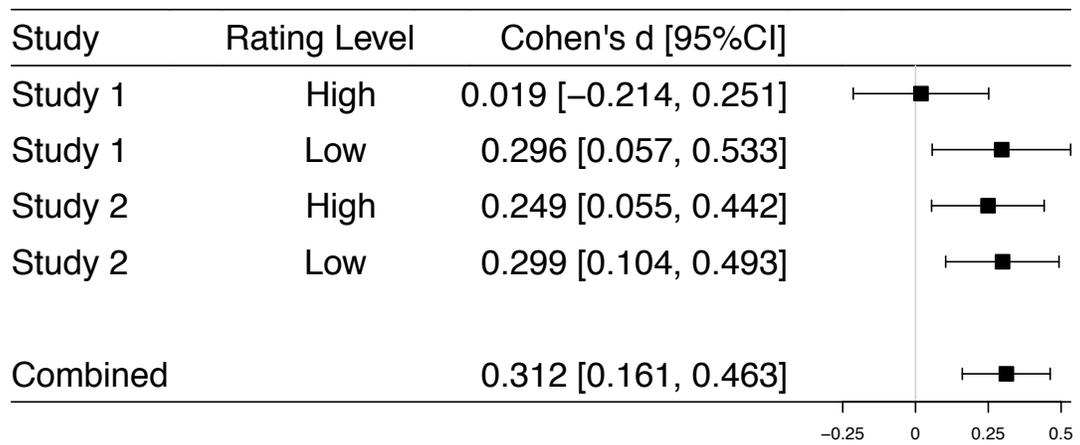| Study	Rating Level	Cohen's d [95%CI]	
Study 1	High	0.019 [−0.214, 0.251]	
Study 1	Low	0.296 [0.057, 0.533]	
Study 2	High	0.249 [0.055, 0.442]	
Study 2	Low	0.299 [0.104, 0.493]	
Combined		0.312 [0.161, 0.463]	

Table 2. Effect Sizes of Go Valuation Effects.

Table 3

Performance on trials with high-rated and low-rated food items in the go/no-go training. Standard deviations are reported in parentheses.

Exp	Parameter	High-Rated	Low-Rated	*t* value	*p*
Exp 1	Go Accuracy	99.2% (1.44%)	98.8% (1.78%)	*t*(70) = 2.31	.024
	No-Go Accuracy	89.0% (11.4%)	90.3% (10.2%)	*t*(70) = -1.11	.269
	Mean GoRT (ms)	393.8 (45.8)	402.1 (45.9)	*t*(70) = -6.66	< .001
	Median GoRT (ms)	390.5 (43.5)	394.5 (42.0)	*t*(70) = -2.78	.007
Exp 2	Go Accuracy	99.4% (1.36%)	98.9% (2.21%)	*t*(105) = 3.37	.001
	No-Go Accuracy	91.1% (9.06%)	92.7% (7.23%)	*t*(105) = -2.13	.036
	Mean GoRT (ms)	410.9 (58.0)	420.0 (57.8)	*t*(105) = -6.42	< .001
	Median GoRT (ms)	405.9 (56.6)	413.7 (55.6)	*t*(105) = -5.53	< .001