1   Does inhibitory control training have an indirect effect on automatic action tendencies for

2   unhealthy foods?

3   Loukia Tzavella[1], Christopher D. Chambers[1], Natalia Lawrence[2], Katherine S. Button[3],

4   Elizabeth Hart[4], Natalie Holmes[4], Kimberley Houghton[4], Nina Badkar[2], Ellie Macey[2],

5   Amy-Jayne Braggins[3], Felicity Murray[2], & Rachel C. Adams[1]

6   [1] Cardiff University Brain Research Imaging Centre, CF24 4HQ, UK

7   [2] School of Psychology, University of Exeter, EX4 4QG, UK

8   [3] Department of Psychology, University of Bath, BS2 7AY, UK

9   [4] School of Psychology, Cardiff University, CF10 3AT, UK

10   Author Note

16   Correspondence concerning this article should be addressed to Loukia Tzavella, Cardiff

17 University Brain Research Imaging Centre, CF24 4HQ, UK. E-mail: tzavellal@cardiff.ac.uk

18 Does inhibitory control training have an indirect effect on automatic action tendencies for

19   unhealthy foods?

Introduction

21   The recent rise in overweight and obesity rates can primarily be ascribed to the

22 over-consumption of energy-dense foods that are high in fat, sugar and salt content (WHO,

23 2018), as individuals are constantly exposed to visual cues of such foods in the environment

24 (e.g., through advertisements) and this often leads to increased food intake (Havermans,

25 2013). A theoretical explanation for this phenomenon has been provided by the dual-process

26 model frameworks which posit that behaviour is determined by the interaction of impulsive

27 (*automatic*) and reflective (*controlled*) cognitive processes (Kakoschke et al., 2015; Strack &

28 Deutsch, 2004). For example, over-consumption of unhealthy foods can be attributed to

29 heightened approach bias for food cues in the environment, which can result in increased

30 food intake if these automatic action tendencies are not regulated via controlled processes,

31 such as inhibitory control (Kakoschke et al., 2017b). Such theoretical frameworks have led to

32 the development of behaviour change interventions for unhealthy eating behaviours that

33 target either automatic or controlled processing, that is approach bias modification and

34 inhibitory control training respectively (see Kakoschke et al., 2017a, and

35 @jones_cognitive_2018 for recent reviews). The primary aim of the present study was to

36 investigate the interaction between automatic and controlled processing in the context of

37 inhibitory control training (ICT). We tested whether strengthening inhibitory control would

38 influence automatic action tendencies towards unhealthy foods after training. To establish

39 whether the employed ICT paradigm was effective we also examined effects of training on

40 impulsive food choice and liking.

41   In the dual-process model frameworks, unhealthy eating behaviours may be explained

42 by a weak reflective system and/or a strong impulsive system (e.g., Lawrence, Hinton,

Parkinson, & Lawrence, 2012; Nederkoorn, Coelho, Guerrieri, Houben, & Jansen, 2012),

which can often be in conflict. For example, automatic attentional (e.g., attending to the

cue) and motivational (e.g., approaching the food) processes would antagonize the controlled

process of considering long-term goals such as losing weight when an individual has to decide

on an action, that is to eat or not eat the food (Kakoschke et al., 2015). In this study we

focus on an automatic process known as approach bias, which is the automatic action

tendency to approach an appetitive (food) cue in the environment, rather than avoid it (C.

E. Wiers et al., 2013). Approach bias has been demonstrated for a variety of appetitive cues,

such as cigarettes (e.g., Bradley, Field, Healy, & Mogg, 2008), alcohol (e.g., Wiers, Rinck,

Kordts, Houben, & Strack, 2010) and cannabis (e.g., Field, Eastwood, Bradley, & Mogg,

2006). In the food domain, there is evidence for the existence of approach bias for a variety

of energy-dense foods (Brignell, Griffiths, Bradley, & Mogg, 2009; Kemps & Tiggemann,

2015; Kemps, Tiggemann, Martin, & Elliott, 2013; Veenstra & de Jong, 2010)[1]. Interestingly,

Kakoschke et al. (2015) found that approach bias alone did not predict increased intake of

unhealthy foods, but it was the interaction between approach bias and inhibitory control

that was the significant determinant of subsequent behaviour. The authors report that

approach bias had the expected effect on food intake only for participants with low

inhibitory control. As an important component of controlled processing, inhibitory control

has been defined as "the ability to inhibit a behavioural impulse in order to attain

higher-order goals, such as weight loss" (Houben, Nederkoorn, & Jansen, 2012, p. 550) and

encompasses several elements, such as response inhibition and cognitive flexibility (see

Bartholdy, Dalton, O'Daly, Campbell, & Schmidt, 2016). Inhibitory control capacity is often

measured via response inhibition paradigms, such as the go/no-go task and stop-signal task,

and has been associated with unhealthy eating behaviours (e.g., Jasinska et al., 2012;

---

[1] There is great variability in terms of methodology for the assessment of both approach bias and inhibitory control in reported studies and therefore the replicability of certain findings is questionable. This and similar issues are presented in the *Discussion.*

Guerrieri et al., 2007; Hall, 2012). Nederkoorn, Houben, Hofmann, Roefs, and Jansen (2010) showed that strong implicit preferences for snacks paired with low "inhibitory control capacity" predicted weight gain over one year.

Complementary evidence for the role of automatic and controlled processes in the regulation of eating behaviours stems from the line of research dedicated to the development of health behaviour change interventions. Approach bias modification training is commonly delivered via an approach-avoidance task (AAT; Neumann & Strack, 2000; Rinck & Becker, 2007; Reinout W Wiers et al., 2013) and has been applied to several unhealthy behaviours involving appetitive cues, such as alcohol consumption and cigarette smoking (e.g., Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011, @wittekind_approach-avoidance_2015). The AAT is assumed to capture automatic action tendencies when participants are instructed to respond to an irrelevant feature of a presented picture, such as the orientation (portrait or landscape), by pulling or pushing a joystick (C. E. Wiers et al., 2013). The AAT can also pair actions with visual feedback, so that the picture gets bigger when participant pull the joystick towards them (zoom-in) and gets smaller when they push it away (zoom-out) . This "zooming" feature can disambiguate the mapping of responses to approach and avoidance actions, whereby pulling the joystick represents approach and pushing it reflects avoidance (Neumann & Strack, 2000). In training, contingencies between actions and stimuli are manipulated so that appetitive cues are associated with push actions (avoidance). Several studies employing AAT variants have found that training can be effective in re-training approach bias for foods (Brockmeyer, Hahn, Reetz, Schmidt, & Friederich, 2015; Kemps et al., 2013) and even reduce food intake in the laboratory (Schumacher, Kemps, & Tiggemann, 2016; for review see Kakoschke et al., 2017a).

In the context of controlled processes, ICT interventions involve food-specific go/no-go or stop-signal tasks whereby participants are instructed to make a speeded choice response to healthy and unhealthy foods, but to withhold that response when a visual, or auditory, signal

is presented. Signal-stimulus mappings are manipulated so that healthy foods are associated with a response (*go* foods) and unhealthy foods are paired with a stop signal (*no-go* foods). Stopping to unhealthy foods has been shown to reduce food consumption (Adams, Lawrence, Verbruggen, & Chambers, 2017; Houben & Jansen, 2011, 2015; N. S. Lawrence et al., 2015; Veling, Aarts, & Papies, 2011; also see Allom, Mullan, & Hagger, 2016 for meta-analysis), promote healthy food choices (Veling et al., 2013a; Veling, Chen, et al., 2017) and has even been associated with increased weight loss (N. S. Lawrence, O'Sullivan, et al., 2015; Veling, van Koningsbruggen, Aarts, & Stroebe, 2014). A potential mechanism of action behind ICT effects on food consumption is stimulus devaluation (Veling et al., 2017), whereby the evaluations of appetitive foods are reduced during training to facilitate performance when response inhibition is required (e.g., Chen, Veling, Dijksterhuis, & Holland, 2016). A possible explanation for this devaluation effect is provided by the Behaviour Stimulus Interaction (BSI) theory which posits that food stimuli are devalued when negative affect is induced to resolve the ongoing conflict between triggered approach reactions to appetitive foods and the need to inhibit responses towards those stimuli (Chen et al., 2016; Veling, Holland, & van Knippenberg, 2008; Veling et al., 2017). When a food is devalued, the approach bias towards that cue is reduced and therefore inhibition can successfully take place. Would it be possible for this reduction in approach bias to be learned via ICT paradigms as well?

In this study we attempt to answer this question by employing a go/no-go training paradigm with unhealthy food stimuli and measure automatic action tendencies via an AAT before and after training to establish whether individuals show reduced approach bias for the foods associated with response inhibition. If an approach action tendency is consistently reduced through devaluation to facilitate inhibition of responses towards appetitive foods, then ICT may have an indirect effect on approach bias. This is also consistent with the concept of an "associative stop system", whereby stimuli associated with stopping can be devalued through an interaction of a stop system and an aversive system (see Verbruggen, Best, Bowditch, Stevens, & McLaren, 2014). Consistent with previous ICT literature, the

study examined impulsive food choice and food liking (i.e., stimulus devaluation

manipulation check) as secondary training outcomes.

Hypotheses

All hypotheses described in this section are confirmatory and have been pre-registered[2]

on the Open Science Framework (https://osf.io/wav8p/). We examined the effects of ICT

(go/no-go training; see *Go/No-go training*) on automatic action tendencies (see *Approach*

avoidance task) and liking (see *Food liking ratings*) for unhealthy foods. These effects were

investigated using change scores from pre-to post-training for both outcomes (H1, H3).

Training condition was also expected to have an effect on food choice behaviour (H2; see

Food choice task). The study assessed contingency learning mechanisms for the training

paradigm, as a manipulation check (H4).

Training effects on automatic action tendencies

The primary outcome measure in the study was the change in automatic action

tendencies from pre-to post- ICT training for the foods associated with different conditions

(go, no-go and control - see Figure 1). Action tendencies were indirectly measured via the

AAT and approach-avoidance bias scores were obtained by subtracting the median response

times (RTs) in avoid trials (push action) from the RTs in approach trials (pull action) at the

participant level, for each training condition and then calculating the change from pre-to

post-training. We hypothesized that ICT training would lead to a reduction in approach bias

for no-go goods and increase in approach bias for go goods compared to the control foods.

---

[2] Exact hypotheses from the pre-registered protocol have been re-ordered according to outcomes for clarity.
We report no deviations from the protocol for the hypotheses and corresponding statistical tests.

H1. There will be *moderate* evidence for an effect of training condition (go, no-go, control) on the change in approach-avoid bias scores from pre-to post-training.

H1a. Participants will show a reduction in approach bias for no-go foods compared to the control foods, from pre-to post- training.

H1b. Participants will have increased approach bias towards go foods relative to the control foods, from pre-to post- training.

Training effects on impulsive food choices

As a secondary outcome, we also examined whether ICT would affect impulsive food choices for unhealthy foods. We compared the probabilities of choosing go and no-go foods compared to control foods. We expected that after ICT participants would show reduced impulsive choices for no-go foods and increased choices for go foods relative to control foods.

H2. Two Bayesian paired samples t-tests were conducted for the mean proportions of selected foods in the go and no-go training condition compared to the control.

H2a. Participants will show reduced choices for no-go foods relative to the control foods.

H2b. Participants will show increased choices for go foods relative to the control foods.

Manipulation check 1: Stimulus devaluation

The mean change in food liking ratings from pre-to post-training was examined for each training condition in order to test whether no-go training led to the devaluation of no-go foods compared to control foods. It should be noted that this was not a positive control for training effectiveness, as the findings for stimulus devaluation outcomes remain controversial (see Jones et al., 2016 for meta-analysis). Stimulus devaluation in this study

161 was therefore treated both as a manipulation check for the employed training paradigm and

162 a secondary outcome measure.

163 H3. There will be *moderate* evidence for an effect of training condition (go, no-go, control)

164 on the change in food liking from pre-to post-training.

165 H3a. Participants will show reduced liking for no-go foods relative to the control foods, from

166 pre-to post- training.

167 H3b. Participants will show increased liking for go foods relative to the control foods, from

168 pre-to post- training.

Manipulation check 2: Contingency learning

170 Training performance was examined in terms of contingency learning. ICT paradigms,

171 such as the go/no-go training task, might lead to stimulus-response associations and learning

172 can be observed in the reaction times and error rates for the different stimulus-response

173 mappings (e.g., N. S. Lawrence, O'Sullivan, et al., 2015). The percentage of successful signal

174 trials (i.e., successful stops) and the reaction times from no-signal (go) trials were compared

175 for specific training conditions, as stated in the hypotheses below.

176 H4. Go/no-go training will result in contingency learning in terms of reaction times on

177 no-signal trials and the percentage of successful inhibitions on signal trials.

178 H4a. Percentage of successful stops will be greater for no-go foods compared to the control

179 foods associated with a signal (control$_{nogo}$).

180 H4b. Go reaction times will be faster for go foods compared to the no-signal control foods

181 (control$_{go}$).

<p style="text-align:center">182   **Methods**</p>

Participants

183

184   255 participants were recruited in total from the University campuses of Cardiff, Bath

185 and Exeter via research participation schemes (e.g., Experimental Management system;

186 EMS) and advertisements (see Figure A1 for recruitment details). Participants recruited

187 through participation schemes received course credits, whereas other individuals were offered

188 entry into a prize draw for one of three £20 shopping vouchers. Participants were informed

189 about the study eligibility criteria and in order to ensure compliance they completed a

190 screening survey in the beginning of the study and provided their consent. They were asked

191 to refrain from eating for 3 hours before the study. Participants had to be at least 18 years

192 of age, be fluent in spoken and written English and have normal or corrected-to-normal

193 vision, including normal colour vision. Participants were excluded if they were dieting at the

194 time of the study, with a weight goal and time-frame in mind, had a current and/or past

195 diagnosis of any eating disorder(s) and had a body-mass-index (BMI) lower than 18.5 kg/m^2

196 (i.e., underweight category). The study was approved by the Ethics Committees of Cardiff

197 University, University of Bath and the University of Exeter.

Sampling plan

198

199   The required sample size was estimated based on a frequentist power analysis

200 conducted for the primary outcome measure (i.e., change in approach-avoidance bias, from

201 pre-to post-training, between go and no-go foods; H1a and H1b) and the stimulus

202 devaluation manipulation check (i.e., change in food liking, from pre-to-post training,

203 between go and no-go foods; H3). Both of these effect sizes were in the medium range, we

204 therefore based our calculations on the primary outcome measure. For an expected effect

205 size we considered other studies that have measured approach bias pre-and

post-approach-avoidance training (Becker, Jostmann, Wiers, & Holland, 2015; Schumacher et al., 2016). Both studies reported an effect size of $\eta_p^2 = 0.07$ which corresponds to a "medium" effect size. Becker et al. (2015) also reported two non-significant results, although effect sizes were not provided. Note, however, that Becker et al. (2015) compared an active group with 90:10 mapping (i.e., avoidance of 90% for unhealthy trials and 10% healthy trials) to a control group with 50:50 mapping whereas Schumacher et al. (2016) compared a 90:10 active group with a 10:90 control group. We therefore took a conservative approach when calculating our sample size. Firstly, we reduced the effect size by 33% (i.e., $dz = 0.34$) to account for publication bias (Button et al., 2013) and secondly we used an alpha of .005, which has recently been recommended for any research that cannot be considered a direct replication and can increase the reliability of new discoveries (Benjamin et al., 2017). Based on a priori power calculations using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) we estimated that a total sample of 149 participants[3] was necessary for 90% power.

The sampling method and power analysis of the study followed a conservative frequentist approach, but the pre-registered analyses were based on a Bayesian framework (see *Pre-registered analysis plan*). Frequentist analyses were also reported in a supplementary fashion ($\alpha = .005$). Bayes factors (BFs) informed the interpretations of the results and although debate exists about labelling evidence in terms of BFs (Richard D. Morey, 2015), we followed the guidelines by (Lee & Wagenmakers, 2013). A threshold of $BF_{10} > 6$ was used to indicate *moderate* evidence for the alternative hypothesis relative to the null, and $BF_{10} < 1/6$ reflected *moderate* evidence for the null relative to the respective alternative hypothesis. Bayes factor analyses were favoured for drawing conclusions from the study, as they would allow us to interpret null outcomes as evidence of absence when traditional analyses would not make such inferences feasible. For frequentist analyses, an alpha level of 0.005 was used.

---

[3] Due to the large number of participant exclusions based on mean error rates in the AAT (see Figure A1) and the group testing laboratory setting at Cardiff University, final recruitment led to the expected sample size including 14 more participants (N=163).

Procedure

The study procedure can be seen in Figure 1 (panel A). After screening, eligible participants were provided with a short survey (see *Survey & Questionnaires*) and proceeded to rate all food categories on how much they like the taste (see *Food liking ratings*). Three blocks of the approach-avoidance task (AAT) were completed before the go/no-go training paradigm was performed. Rated food categories were randomly assigned to three conditions for training: go, no-go and control, as shown in Figure 1 (panel B). Post-training, participants were presented with another three blocks of the AAT, provided ratings for all food stimuli again and finally completed a short food choice task (see *Food choice task*). At the end of the study, several questionnaires were presented in random order (see *Survey & Questionnaires*) and participants were debriefed about the aims of the study.
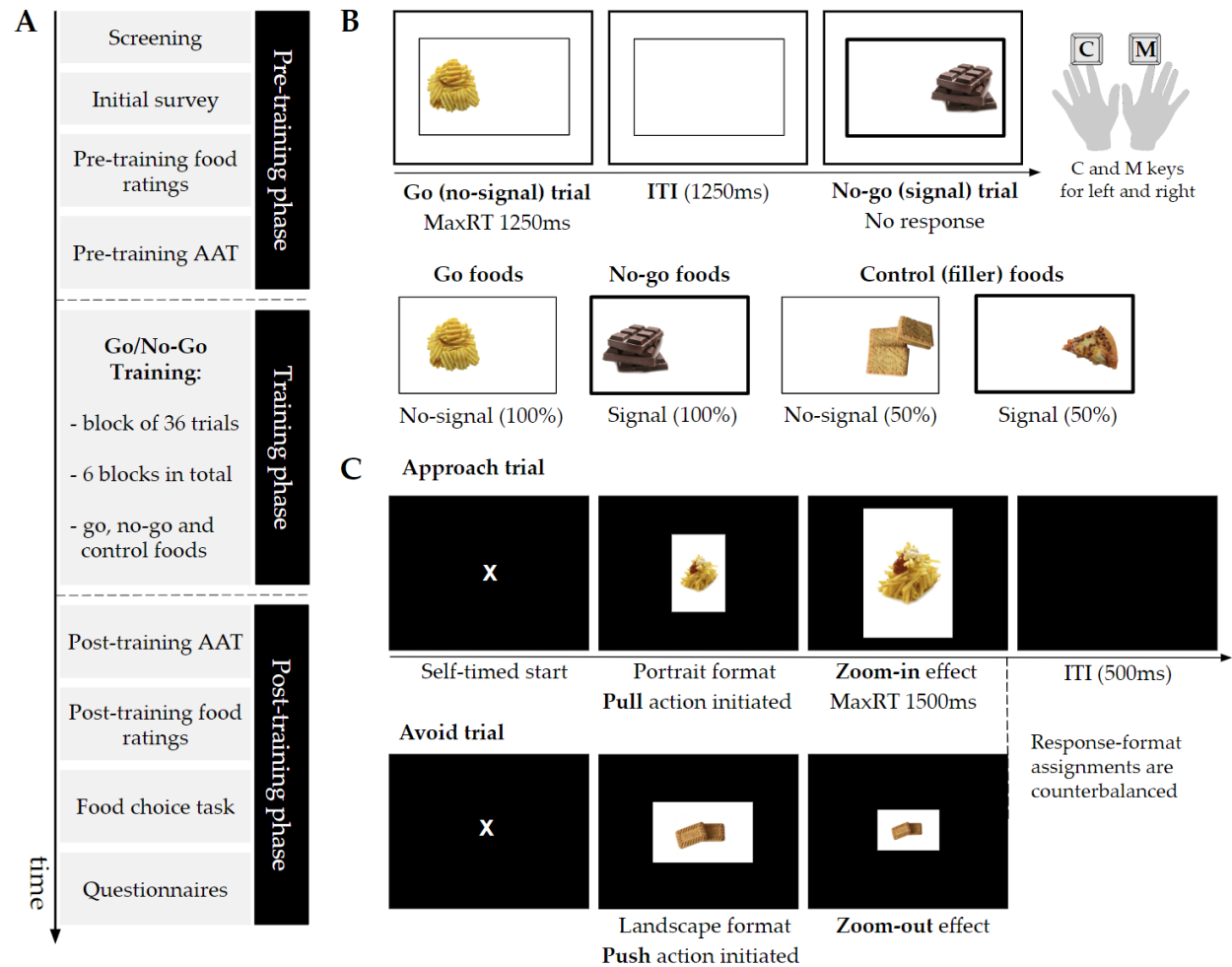
Figure 1. **Schematic diagram of the study procedure, go/no-go training and approach-avoidance tasks. A.** After completing the screening and initial survey, participants rated all food stimuli (liking) and proceeded to perform the pre-training approach-avoidance task (AAT) blocks. In the training phase, participants completed six blocks of go/no-go training. The post-training AAT blocks were then presented and followed by food liking ratings. At the end of the study, participants completed a short food choice task and several questionnaires, in random order. **B.** The go/no-go training paradigm involved go (no-signal) and no-go (signal) trials that occurred with equal probability. On go trials, participants had to respond within 1250ms by pressing the "C" and "M" keys to indicate the picture location (left or right, respectively). On no-go trials, participants were instructed not to respond at all. The inter-trial interval (ITI) was 1250ms. Food categories were randomly assigned to three conditions. Go foods were only paired with no-signal trials and no-go foods were always associated with no-signal trials. Control, or filler, foods were presented in both signal and no-signal trials (50:50).

C. In the AAT, participants were asked to respond according to the format of the presented rectangle (portrait or landscape). Response-format assignments were approximately counterbalanced across participants. As an example, on approach trials a participant would have to pull the mouse towards them when the picture was in portrait format (approach trial) and push it away from them when the picture was in landscape format. Push and pull actions were paired with visual feedback, that is, zoom-out and zoom-in effects respectively. The maximum reaction time (maxRT) was 1500ms and the ITI was set to 500ms. Participants clicked on a central "X" to begin a trial (self-timed start).

Go/No-go training

The Go/No-Go (GNG) training paradigm involved go and no-go responses to six pre-selected appetitive food categories. Food categories differed in terms of taste, so that three foods were savoury (i.e., pizza, crisps, chips) and three foods were sweet (i.e., biscuits, chocolate, cake)[4]. Two food categories were randomly assigned to each training condition (go, no-go, filler foods) in the beginning of the experiment and food taste was counterbalanced so that each condition had one sweet and one savoury food. There were three training conditions according to the mapping of foods to signal (no-go) and no-signal (go) trials in the GNG. All go foods appeared in go (no-signal) trials and all no-go foods were presented in no-go (signal) trials (see Figure 1, panel C). Control, or filler, foods appeared on both go and no-go trials with equal probability (i.e., 50:50). Each food category had three exemplars which appeared twice in each block.

All foods were presented on either the left or right hand side of the screen within a rectangle for 1250ms, which was the maximum reaction time (maxRT), as shown in Figure 1, panel B. Participants were asked to respond to the location of the food as quickly and as accurately as possible by pressing the "C" and "M" buttons on the keyboard with their left and right index fingers, respectively. The central rectangle remained on the screen

─────

[4] All study materials are openly available at https://osf.io/wcf4r/

throughout the training, including the inter-trial-interval (ITI), which was 1250ms. On signal trials, the rectangle turned bold, indicating that participants should withhold their response. In line with the GNG training paradigm, this signal appeared on stimulus onset (i.e., no delay between stimulus and signal) and stayed on the screen until the end of the trial. A correct response on no-signal trials was registered when participants responded accurately to the location of the food within the maxRT window and a successful stop (i.e., correct signal trial) was considered when participants did not respond at all. Incorrect responses in no-signal trials refer to either to a wrong location judgement or a missed response. Left and right responses were counterbalanced across all manipulated variables for each type of trial. Training was split into 6 blocks of 36 trials (216 trials in total) and lasted approximately 10 minutes with inter-block breaks (15s). Task practice included 12 trials of go and no-go responses (50%-50%) and participants responded to the location of grey squares, instead of food pictures. For the practice trials, we provided accuracy feedback during the ITI.

Approach avoidance task

The approach-avoidance task (AAT) was adapted from an existent paradigm (Rinck & Becker, 2007; Wiers, Rinck, Dictus, & Van Den Wildenberg, 2009), which involves "pull" (i.e., towards self) and "push" (i.e, away from self) movements of a joystick. Each type of motor response is paired with visual feedback so that when the joystick is pulled, the image gets bigger (zoom-in) and when it is pushed, the image gets smaller (zoom-out). This "zooming" effects acts as an exteroceptive cue of either an approach or avoidance response (Neumann & Strack, 2000). This feature of the joystick AAT complements the proprioceptive properties of the task, where responses requiring arm flexion and extension correspond to approach and avoidance trials, respectively. This task also disambiguates approach and avoidance responses by using the "zooming" feature (Wiers et al., 2009). For example, arm extension could indicate an approach response towards an appetitive food (object-reference)

or an avoidance response where the food is pushed away from the body/self (self-reference; Phaf, Mohr, Rotteveel, & Wicherts, 2014). The visual feedback thus provides the self-reference attribute to the responses (e.g., object comes closer to one's body). We also adopted the evaluation-irrelevant feature of the paradigm, whereby participants respond according to the format of the picture (portrait or landscape; e.g., Wiers et al., 2010).

AAT responses involved "push" and "pull" movements of the computer mouse (adaptation of the joystick version). Food stimuli were presented in the centre of the screen and participants were instructed to pull the mouse towards them or push the mouse away from them according to whether the image was in portrait or landscape format (see Figure 1, panel C). Response-format assignments were approximately counterbalanced across participants (45.4% portrait-approach, 54.6% landscape-approach). Instructions highlighted moving the mouse cursor until it reaches the end of the screen (top or bottom edge) for a correct response to be registered and making smooth whole-arm movements. Participants had 1500ms to respond after the stimulus appeared. Each trial started with a central "X" on the screen and participants had to click on it to begin (self-timed start). The ITI was 500 ms and there was no delay between the "X" click response and the stimulus onset. In order to account for the natural movement of the mouse, pixel tolerance was added to every mouse movement ($\pm$ 1.25% of display height), including movement initiation in the beginning of the trial. A response in the AAT was registered as correct only when participants completed the correct action (e.g., pull or push) within the maxRT window and also initiated a movement towards the correct direction. Even if the final response was correct, participants could have changed their movement after making an initial error (e.g., pull instead of push the mouse in an "avoid" trial) and therefore the direction of their initial movement was also taken into account. The complete RT for an AAT trial was defined as the time from the stimulus onset to the successful completion of a response.

Each AAT block consisted of 72 trials and go, no-go and control foods appeared with

309   equal probability for both "pull" (approach) and "push" (avoid) responses. There were 12

310   approach and 12 avoid trials for each training condition (e.g., no-go) and within those trials,

311   there were six savoury and six sweet foods presented (i.e., three exemplars repeated twice).

312   Three AAT blocks were performed before training (AAT$_{pre}$) and three after training

313   (AAT$_{post}$). There was a number of constraints placed on the quasi-random order of the trials

314   within an AAT block. There were no more than three images of the same food category

315   being presented consecutively and no more than three trials with the same picture format in

316   sequence. AAT practice consisted of 10 trials with grey rectangles instead of food stimuli

317   and accuracy feedback. The screen background for the AAT was black and the task lasted

318   approximately 15 minutes, including the inter-block 15s breaks, where participants received a

319   reminder of the main instructions.

Food liking ratings

321   Participants provided food liking ratings before and after training using a visual

322   analogue scale (VAS). They rated all foods included in the GNG paradigm according to how

323   much they liked the taste, ranging from 0 ("not at all") to 100 ("very much"). Task

324   instructions encouraged participants to imagine they were tasting the food in their mouth

325   and then rate how much they liked the taste. The order of the presented foods was

326   randomised and each block consisted of 18 trials. Participants completed a block before

327   training (Liking$_{pre}$) and a block after training (Liking$_{post}$).

Food choice task

329   Impulsive food choices were assessed using a food choice task adapted from Veling et al.

330   (2013a), which included all food categories from the GNG paradigm (two exemplars per

331   category). The twelve foods were presented on a grid layout and participants had ten

seconds to select three foods that they would like to consume the most at that specific time, by clicking on them with the computer mouse. Participants were asked to click on a "start" button to begin the trial and when a response was registered the selected food stimulus disappeared from the screen. We assumed that this task element would prevent participants from deliberating on their choices and changing their initial responses, which would mean that *impulsive* food choices were no longer measured. However, it should be noted that although participants were not informed about the hypothetical nature of their choices, it is highly probable that they would not consider their choices consequential (i.e., they would not think they would get a food item at the end of the study).

Survey & Questionnaires

Eligible participants were presented with an initial survey to record demographics and other variables for exploratory analyses. The survey consisted of height and weight measurements to calculate participant's body-mass-index (BMI; kg/m^2), the number of hours since their last meal ("less than 3 hours ago", "3-5 hours ago", "5-10 hours ago", "more than 10 hours ago") and hunger state at the the time of the study (VAS: 1="Not at all" to 9="Very"). Gender was also recorded with the options of male, female, transgender male, transgender female, gender variant/non-conforming, and an open ended text response for "other".

Several questionnaires were completed by the participants at the end of the study for exploratory analyses, as part of the undergraduate student projects of the GW4 Undergraduate Psychology Consortium 2017/2018. The Barratt Impulsivity Scale (BIS-15; Spinella, 2007) was introduced as a measure of impulsivity and the Stop Control Scale (SCS; De Boer, van Hooft, & Bakker, 2011) was used to examine a distinctive element of general trait self-control, referred to as stop control. Other administered questionnaires included the Food Cravings Questionnaire - Trait - reduced (FCQ-T-r; Meule, Hermann, & Kübler, 2014),

357 Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983) and the "food" and

358 "money" subscales from the Delaying Gratification Inventory (DGI; Hoerger, Quirk, & Weed,

359 2011). A correlation matrix of main questionnaire measures and sample characteristics can

360 be found in Appendix B.

Analyses

Measures & indices

363 The mean error rates in no-signal and signal trials as well as mean reaction time in

364 no-signal trials (GoRT) from the GNG informed participant exclusions (see *Data exclusions*).

365 For the contingency learning manipulation check (H3, H4), we measured the average

366 proportion of successful stops from signal trials for no-go foods and control foods which were

367 paired with a signal (control-nogo) and the mean GoRTs for each participant from go and

368 control-go trials.

369 Performance in the AAT_{pre} and AAT_{post} blocks was considered only for correct

370 responses. We calculated median RTs for "push" and "pull" responses on all training

371 condition levels, at a participant level[5]. Medians were used instead of means as they are less

372 sensitive to outliers in RT distributions and in line with previous literature (Wiers et al.,

373 2009, 2010). The approach-avoid bias score for each condition was calculated as the

374 difference between the median RTs for "push" and' pull' responses ($MedianRT_{push}$-

375 $MedianRT_{pull}$). Bias scores were computed for both AAT_{pre} and AAT_{post} blocks. Positive

---

[5] RTs were recorded continuously from movement initiation to response completion with samples every 33ms (two display refresh rates) to allow dynamic zoom-in/zoom-out effects based on participants' mouse movements. However, a bug was encountered with the version of the software and the temporal resolution at which coordinates and times were recorded was reduced. For this reason, we used linear interpolation to increase our samples to 100 for every trial and obtain more precise RT measures. All details regarding this procedure and the software bug can be found in our analyses scripts.

scores indicate an approach bias towards the foods of interest and negative scores reflect avoidance for those foods. Change scores for approach-avoid biases from pre-to post-training (ΔAAT bias score) were calculated for pre-registered analyses (H1). The proportion of correct responses for each AAT design cell informed participant exclusions.

Participants were required to choose three foods out of twelve in the food choice task and selections could vary in their number for each training condition (go, no-go, control). Food choices were therefore normalised according to the total number of responses per participant (i.e., proportion). For the analyses of food choices, we compared the mean proportions of choices (calculated per participant relative to their total number of choices) between each training condition. Food rating VAS scores were averaged (mean) across the two foods per training condition (i.e., sweet and savoury foods for go, no-go and control conditions) and the three exemplars of each food. Changes in food liking were examined in terms of change scores (ΔFood liking score) from pre-to-post training.

Data exclusions

Participant-level data exclusions were conducted based on GNG training and AAT performance. Participants who met any of the following criteria were excluded from all respective analyses. We excluded participants who had a mean GoRT greater than three standard deviations from the group mean and percentage of correct responses in no-signal trials less than 85%. Participants were also excluded if their percentage of errors in signal trials was greater than three standard deviations from the group mean and percentage of errors in either pre- or post- AAT blocks greater than 0.25. Additionally, participants who submitted a food rating of 50 (i.e., neutral) for 24 or more trials wither pre-or post-training would not be included as we assumed that multiple such responses would indicate that participants used the default setting of the VAS and purposefully skipped the rating trials.

Pre-registered analysis plan

Data pre-processing and analyses were conducted in RStudio (RStudio Team, 2016) and JASP (JASP Team, 2018). Pre-registered analyses are described under their pre-specified hypotheses, as presented in *Hypotheses*. For all Bayesian paired samples t-tests mentioned hereinafter, we used a prior with the $\sqrt{2/2}$ scale parameter for the half-Cauchy distribution.

H1. The effect of training condition on the change in approach-avoid bias scores from pre-to post-training was examined using a Bayesian Repeated Measures ANOVA with the default prior settings (Rouder, Engelhardt, McCabe, & Morey, 2016; Rouder, Morey, Speckman, & Province, 2012) and participants treated as a nuisance term.

H1a. $\Delta\text{AAT}_{\text{nogo}} < \Delta\text{AAT}_{\text{control}}$

H1b. $\Delta\text{AAT}_{\text{go}} > \Delta\text{AAT}_{\text{control}}$

H2. Two Bayesian paired samples t-tests were conducted for the mean proportions of selected foods in the go and no-go training condition compared to the control.

H2a. p(no-go) < p(control)

H2b. p(go) > p(control)

H3. The effect of training condition on the change in food liking from pre-to post-training was examined using a Bayesian Repeated Measures ANOVA, consistent with H1.

H3a. $\Delta\text{Liking}_{\text{nogo}} < \Delta\text{Liking}_{\text{control}}$

H3b. $\Delta\text{Liking}_{\text{go}} > \Delta\text{Liking}_{\text{control}}$

H4. Contingency learning during go/no-go training was examined using Bayesian paired-samples t-tests for the percentage of successful inhibition trials and go reaction times.

422 H4a. PCsignal$_{\text{nogo}}$ > PCsignal$_{\text{control-nogo}}$

423 H4b. GoRT$_{\text{go}}$ < GoRT$_{\text{control-go}}$

424   The evidential value of confirmatory findings was solely determined by the Bayesian

425 tests outlined in this section, as previously explained (see *Sampling plan.* Frequentist tests

426 were conducted in order to further the reproducibility of findings (e.g., potential use in

427 meta-analyses). Paired samples t-tests were two-tailed, in line with the reported power

428 analysis\footnote{Although we pre-registered Bonferroni corrections for paired sample

429 t-tests following Bayesian Repeated Measures ANOVAs, since we use a very conservative

430 alpha level and have only two contrasts of interest per hypothesis, we did not apply such

431 corrections in the reported p-values. Assumptions for repeated measures ANOVA were

432 checked in line with the pre-registered analysis plan and there were no violations. However,

433 contingency plans were not considered in case the normality assumption was violated for

434 paired t-tests, but appropriate tests are reported in the *Robust statistics* section.

Results

Sample characteristics

437   The final sample for pre-registered analyses consisted of 163 participants (80.98%

438 female). Detailed participant-level exclusions are presented in Figure A1. Participants had

439 on average a healthy BMI (M = 22.88, SD = 2.98, range = 18.54 - 32.36) and their mean

440 age was 22.39 (SD = 9.04, range = 18-59). 108 participants (66.26%) reported that they had

441 their last meal 3-5 hours before the study and hunger levels at the beginning of the study

442 were not particularly high (M = 5.70, SD = 2.22). However, 24 participants (14.72%) did

443 not adhere to the instruction not to eat three hours before the study, as they reported

444 having their last meal "less than 3 hours ago".

Confirmatory findings

Training outcomes. We found *strong* evidence for the absence of a general effect of go/no-go training condition on the change in approach-avoidance bias scores $[BF_{01} = 16.06;$ $F(2, 324) = 1.01, p = 0.365]$. Results for paired comparisons are shown in Table 1. There was *moderate* evidence that the change in bias scores for no-go foods ($\Delta \text{AAT}_{\text{nogo}}$; $M = $ -3.31, $SD = 62.91$) was not reduced compared to the change for filler foods ($\Delta \text{AAT}_{\text{control}}$; $M = $ -1.81, $SD = 59.55$). Similar to H1a, we found *strong* evidence for the null compared to the alternative for H1b. The change in bias scores for go foods ($\Delta \text{AAT}_{\text{go}}$; $M = $ -10.47, $SD = $ 59.57) was not greater than the change for filler foods. Approach-avoidance bias scores pre- and post- training across training conditions can be visualised using rainclouds in Figure 2. The raincloud plots have been created using existing code and guidelines (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2019, 2018).

The effect of training on impulsive food choices was examined for no-go and go foods compared to control, as stated in H2a and H2b respectively. There was *extreme* evidence that the probability of choosing a no-go food ($M = 0.21$, $SD = 0.27$) was reduced compared to the probability of choosing a filler food ($M = 0.36$; $SD = 0.31$) after training[6] (see Table 1). In contrast, we only obtained *anecdotal* evidence that probability of choosing a go food ($M = $ 0.44; $SD = 0.33$) was not greater than the probability of choosing a filler food after training.

Manipulation checks for training. As a first manipulation check for training outcomes, we investigated whether GNG changed the evaluations of foods associated with signal and no-signal trials compared to the evaluations of filler foods which were paired with either type of trial with equal probability (control). We only found *anecdotal* evidence for the absence of a general effect of training condition on the changes in liking from pre- to post- training [H3; $BF_{01} = 2.89$; $F(2, 324) = 2.90$, $p = 0.057$]. The change in liking scores

───────

[6] There was a missing value for this analysis as one participant did not complete the food choice task.

Table 1

Results for all pre-registered hypotheses and respective statistical tests

	BF_{10}	t	df	p	d	95% CI for d		Evidence interpretation
						Lower	Upper	
H1a	0.107	-0.25	162	0.805	-0.02	-0.17	0.13	*Moderate* evidence for H_0
H1b	0.039	-1.35	162	0.179	-0.11	-0.26	0.05	*Strong* evidence for H_0
H2a	247.782	-3.93	161	< .001	-0.31	-0.47	-0.15	*Extreme* evidence for H_1
H2b	0.849	1.82	161	0.070	0.14	-0.01	0.30	*Anecdotal* evidence for H_0
H3a	2.648	-2.38	162	0.019	-0.19	-0.34	-0.03	*Anecdotal* evidence for H_1
H3b	0.067	-0.37	162	0.715	-0.03	-0.18	0.13	*Strong* evidence for H_0
H4a	140.254	3.77	162	< .001	0.30	0.14	0.45	*Extreme* evidence for H_1
H4b	3973.214	-4.66	162	< .001	-0.37	-0.52	-0.21	*Extreme* evidence for H_1

Note. Evidence is interpreted for the alternative hypothesis (H_1) compared to the null (H_0) and vice versa. All Bayesian paired samples t-tests were directional, as indicated in the *Pre-registered analysis plan* section and frequentist equivalents were non-directional (two-tailed). The effect size is represented by Cohen's *d*.

from pre-to post-training for nogo foods ($\Delta\text{Liking}_{\text{nogo}}$; M = -4.16; SD = 9.51) was only slightly reduced compared to change in liking for filler foods ($\Delta\text{Liking}_{\text{control}}$; M = -2.61, SD = 8.77), and there was only *anecdotal* evidence for this effect (H3a; see Table 1). The change in liking scores from pre-to post-training for go foods ($\Delta\text{Liking}_{\text{go}}$; M = -2.87, SD = 10.15), however, was not greater than the change for filler foods as originally expected. Instead, there was *strong* evidence for the null hypothesis compared to the alternative (H3b).

In order to validate whether the implemented go/no-go training paradigm led to stimulus-response associations (i.e., contingency learning) we first tested whether the percentage of correct responses for no-go foods (i.e., successful inhibitions) would be greater compared to the percentage of correct responses for filler foods associated with signal trials (H4a). There was *extreme* evidence that participants had on average more successful
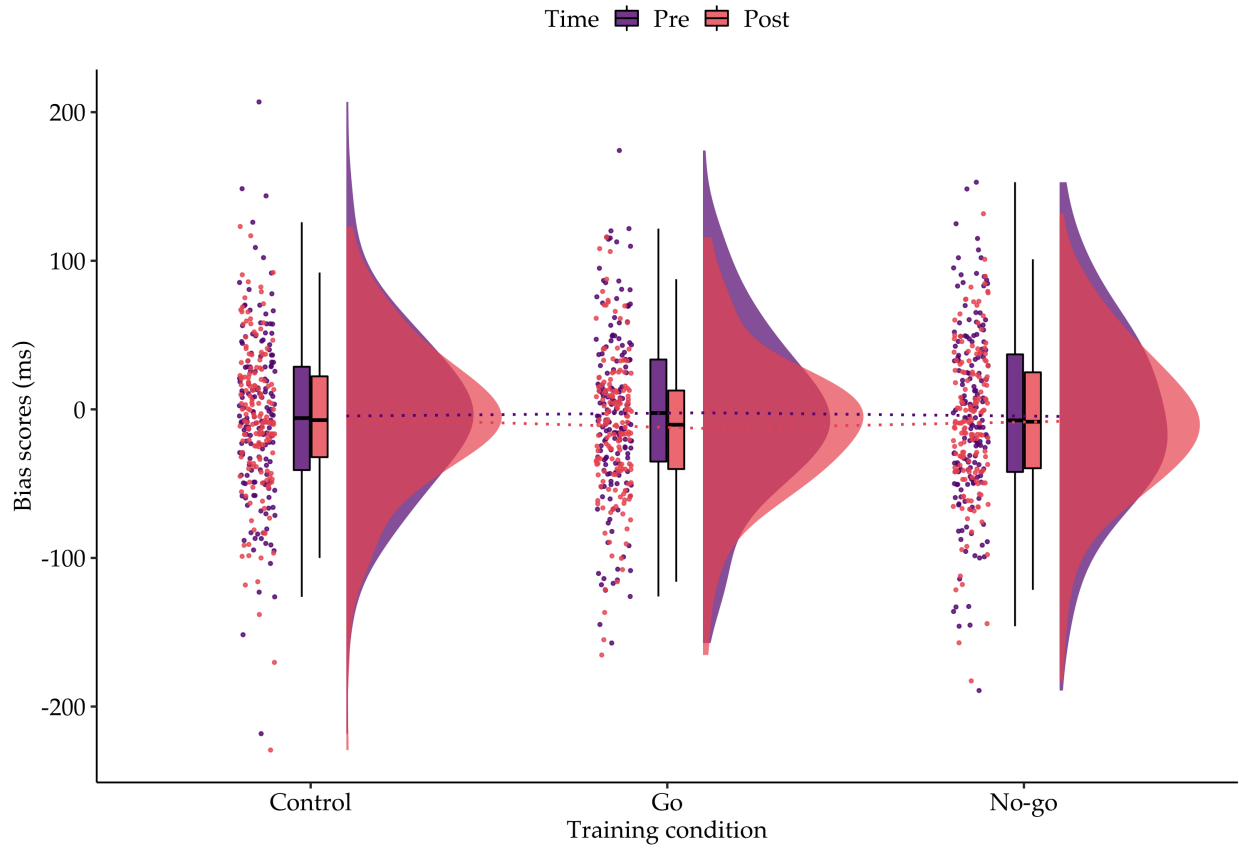
Figure 2. **Raincloud plot of the approach-avoidance bias scores pre- and post- training across training conditions.** There were no differences between the sample mean changes in approach-avoidance bias scores for no-go and go foods compared to control (filler) foods, as shown by the dashed lines. At a closer inspection, individual bias scores do not seem to be clustered around the positive end of the distribution as it would be expected for for appetitive unhealthy foods, but actually around zero and exploratory analyses showed that baseline bias scores did not statistically deviate from zero (see *Baseline approach bias scores*). Also, on the negative end of the distributions the data points show greater dispersion and hint to the potential existence of several outliers. Therefore, we also conducted a robustness check for H1 by employing the most common outlier removal techniques in studies that utilise the AAT (see Section X). A last trend in the data displayed in the rainclouds is the possible difference in bias scores from pre- to post-training for go foods alone (Post < Pre). *Note.* The 'split-half violin' elements in the plot show smoothed distributions.

480 inhibitions for no-go foods (PCsignal$_{nogo}$; $M = 0.97$, $SD = 0.03$) than filler foods

481 (PCsignal$_{control-nogo}$; $M = 0.96$, $SD = 0.04$). For H4b, we tested whether mean reaction times

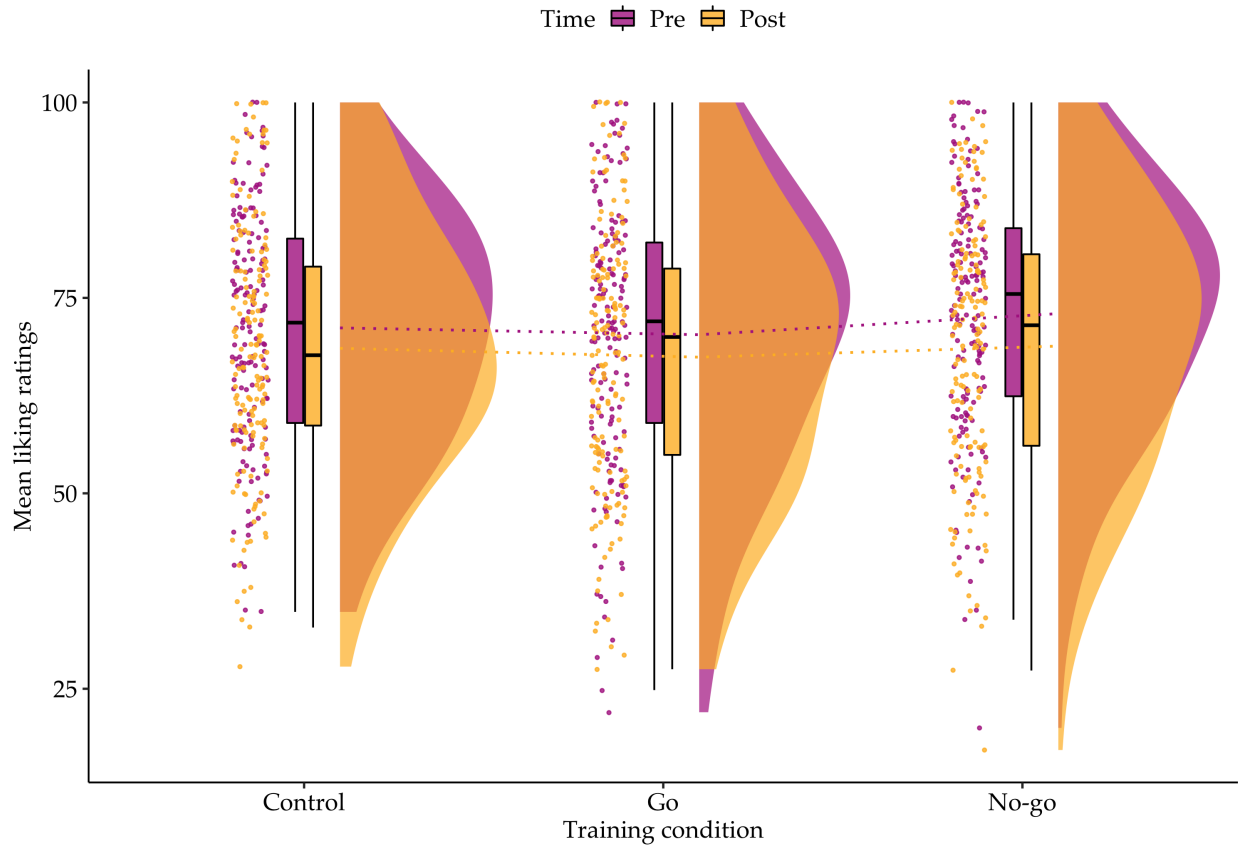482 would be reduced for go foods (GoRT$_{go}$; $M = 507.00$, $SD = 70.48$) compared to filler foods

Figure 3. **Raincloud plot of the mean liking ratings pre- and post- training across training conditions.** This visualisation of the mean liking ratings from all participants revealed that (1) there is a general trend of devaluation across training conditions and (2) the distributions are more right-skewed than expected, towards the least liked range of the visual analogue scale (VAS). As a reminder, taste (liking) ratings were registered on a VAS ranging from 0 to 100 (i.e., 50=neutral). Although there appears to be a small difference between the change in liking for no-go foods compared to the control, the trends presented in this plot should be inspected further to establish whether any findings related to H3 are robust. The presence of potential outliers and the fact that go foods seem to be rated more negatively after training (Post) relative to before (Pre) also informed exploratory analyses presented in Section X. *Note.* The 'split-half violin' elements in the plot show smoothed and trimmed distributions.

483   associated with no-signal trials (GoRT$_{\text{control-go}}$; $M = 515.00$, $SD = 75.51$) and there was

484   *extreme* evidence for such an effect. Therefore, contingency learning was observed in the

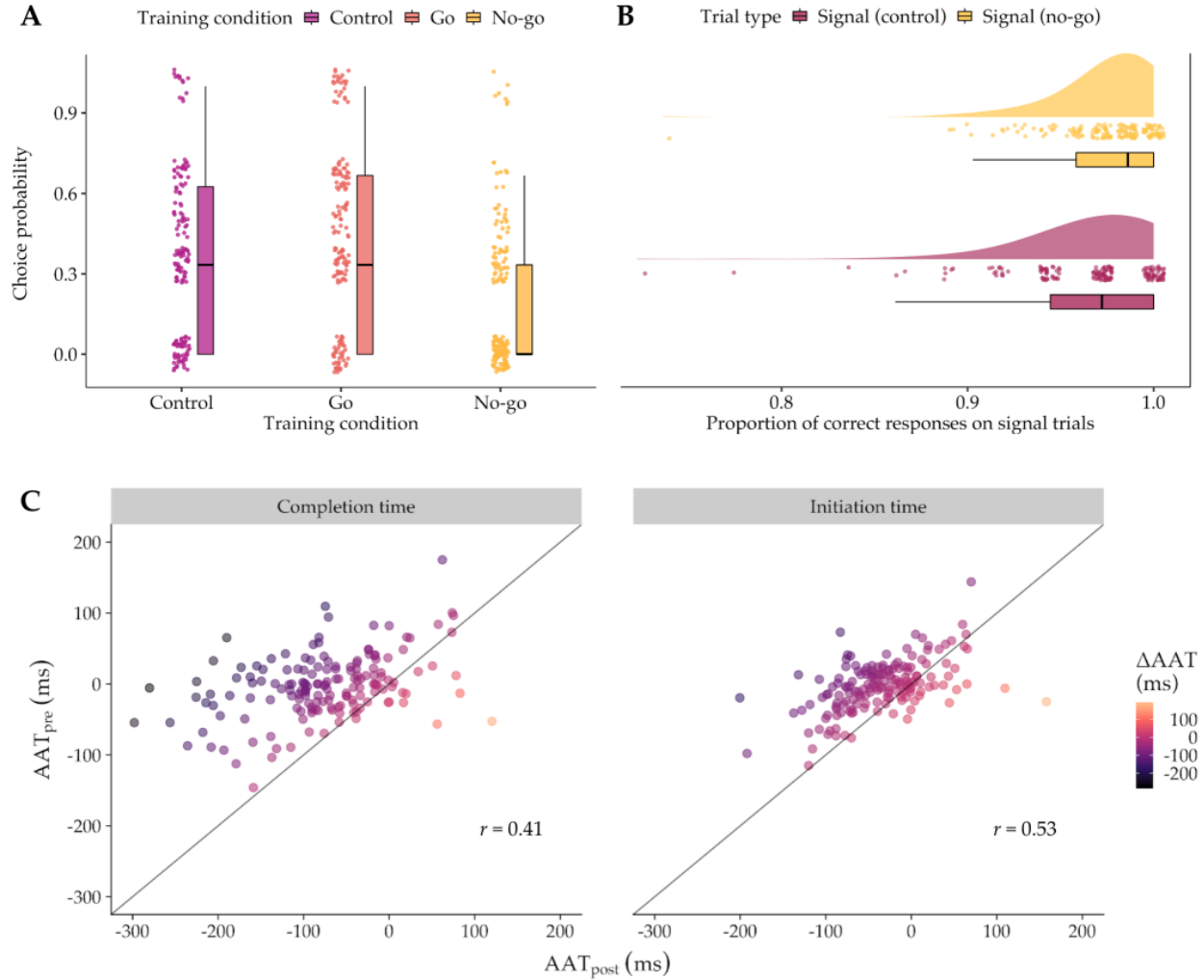485   employed GNG paradigm for both reaction time and accuracy outcomes.

Figure 4. **Plots for food choice outcomes, accuracy-based contingency learning in training and test-retest reliability of approach-avoidance bias scores. A.** The boxplots with corresponding jittered individual data points clearly show that the probability of choosing a no-go food after training is reduced compared to the probability of choosing control food [H2a]. Contrary to our predictions, the average choice probability was not greater for go foods relative to the control [H2b]. **B.** Although the pre-registered statistical analyses for accuracy-based contingency learning in the go/no-go training paradigm [H4a] suggest that the proportion of correct responses on signal trials (PC_{signal}) for would be on average greater for no-go foods compared to control foods, as shown by the drawn means (diamond symbols) on the boxplot, we observe an effect in the opposite direction with the medians. It is not surprising that the accuracy distributions are heavily skewed (normality violated), but we suggest that findings regarding this learning effect are interpreted with caution. Learning would also be expected to manifest as slight differences between the two halves of the task blocks, but the distributions are almost identical between the first and second half. Results from the same analysis, after exclusion of outliers, are reported in the main text.

C. For the pre-registered analysis of training outcomes on automatic action tendencies [H1], approach-avoidance bias scores (ΔAAT) were calculated based on median reaction times for pull and push responses from pre-to post-training (AAT$_{\text{post}}$ - AAT$_{\text{pre}}$). These reaction times refer to the time participants took to complete an action (i.e., completion time). The test-retest reliability of these scores is Section X. *Note.* The 'split-half violin' elements in the plot show smoothed and trimmed distributions.

Exploratory findings

Baseline approach bias scores. Performance in the AAT was inspected further to check if approach bias for foods was present in the final sample and whether error rates differed across conditions. Although the sample means for AAT$_{\text{pre}}$ bias scores are negative for go foods (M = -2.32, SD = 58.14), no-go foods (M = -4.75, SD = 60.58) and filler foods (M = -4.48, SD = 52.25), individual data points (see Figure 2 show less dispersion close to zero, suggesting that neither approach or avoidance bias was captured by the AAT. In line with previous literature (see Table 1 in Becker et al., 2015), we directly tested this hypothesis by examining whether baseline bias scores statistically deviated from zero using Bayesian one sample t-tests with the default prior settings for the two-sided alternative hypothesis that the population mean was larger than the test value. Equivalent frequentist tests were also conducted. Overall, we obtained conclusive evidence for the absence of baseline approach/avoidance bias for any of the foods which were randomly assigned to training conditions (see Table 2).

As baseline bias scores calculated from completion times may be "contaminated" by motor demands in this version of the AAT that requires computer mouse movements and arm flexion/extension, we considered the possibility that motor initiation times may be more sensitive to capturing automatic action tendencies. Movement initiation was registered when participants had moved their mouse cursor since starting a trial (i.e., stimulus onset), including the pixel tolerance for natural movements of the mouse (see *Approach avoidance*

506 *task*). We therefo re checked whether baseline bias scores calculated using median initiation

507 times, instead of median completion times, statistically deviated from zero. Consistent with

508 the results presented above, we found strong evidence that they did not deviate from zero

509 across training conditions (see Table 2)

Table 2

Results of Bayesian and frequentist one sample t-tests for baseline approach-avoidance bias scores

					95% CI for d	
	BF_{01}	$t(162)$	p	d	Lower	Upper
Completion time: AAT_{pre} for go foods	10.08	-0.51	0.611	-0.04	-0.19	0.11
Completion time: AAT_{pre} for no-go foods	7.01	-1.00	0.318	-0.08	-0.23	0.08
Completion time: AAT_{pre} for control foods	7.02	-1.00	0.319	-0.08	-0.23	0.08
Initiation time: AAT_{pre} for go foods	10.73	-0.36	0.718	-0.03	-0.18	0.13
Initiation time: AAT_{pre} for no-go foods	10.18	-0.49	0.626	-0.04	-0.19	0.12
Initiation time: AAT_{pre} for control foods	10.46	-0.43	0.669	-0.03	-0.19	0.12

Note. AAT_{pre}: Pre-training approach avoidance task bias scores

510   **Sub-group analysis.**   In an effort to show that training did not have an effect on

511 AAT outcomes was not due to the absence of baseline approach bias for unhealthy foods, we

512 conducted a sub-group analysis for participants with positive baseline bias scores (N=72).

513 There was *very strong* evidence for the absence of a main effect of go/no-go training

514 condition on the change in approach-avoidance bias scores [$BF_{01} = 43.99$; $F(2, 142) = 0.01$,

515 $p = 0.987$]. For this sub-group food-choice outcomes were consistent with the confirmatory

516 findings reported in *Training outcomes*. There was *strong* evidence that the probability of

517 choosing a no-go food ($M = 0.20$, $SD = 0.27$) was reduced relative to the probability of

518 choosing a control food ($M = 0.37$, $SD = 0.33$) [$BF_{10} = 13.97$, $t(70)] = -2.96$, $p = 0.004$, d

519 $= -0.35$, 95% CI for $d = -0.59, -0.11$]. There was *moderate* evidence for the absence of a

520 general effect of training condition on the change in liking scores from pre-to post-training

521 $[BF_{01} = 8.91; F(2, 142) = 0.94, p = 0.392]$. With regards to the contingency learning

522 manipulation check, there was *very strong* evidence for a greater proportion of correct

523 responses in signal trials with no-go foods compared to control foods $[BF_{10} = 37.80, t(71) =$

524 3.33, $p = 0.001$, $d = 0.39$, 95% CI for $d = 0.15$, 0.63]. However, there was only *anecdotal*

525 evidence that GoRTs were faster for go foods compared to control foods $[BF_{10} = 3.52, t(71))$

526 $= -2.38$, $p = 0.020$, $d = -0.28$, 95% CI for $d = -0.52$, -0.04].

527 **Accuracy in the approach-avoidance task.**   Although reaction times are the

528 primary measure of interest for studies that utilise the AAT so that bias scores can be

529 calculated, we consider essential to also report an exploratory examination of error rates. At

530 baseline, average error rates were not increased for trials where participants were required to

531 avoid an appetitive food and complete a push action ($M = .136$; $SD = .070$) relative to trials

532 where an approach (pull) action was completed ($M = .143$; $SD = .066$) $[BF_{01} = 26.49$;

533 $t(162) = 1.42$, $p = 0.159$, $d = 0.11$, 95% CI for $d = -0.04$, 0.26]. However, after training

534 participants had on average more errors in approach trials ($M = .124$, $SD = .074$) compared

535 to avoid trials ($M = .105$, $SD = .062$) $[BF_{10} = 90.98$, $t(162) = 3.64$, $p < .001$, $d = 0.29$, 95%

536 CI for $d = 0.13$, 0.44]. It is possible that training had a "hidden" effect on accuracy, for

537 example it was more difficult to approach no-go foods compared to go and/or control foods

538 due to a learned association between response inhibition and these food stimuli. We

539 calculated difference scores from the mean error rates post-training (pull - push) to check

540 whether this increase in error rates was general or specific to training conditions. There was

541 very strong evidence for the absence of a general effect of training condition on differences in

542 mean error rates between approach and avoid trials H3; $BF_{01} = 37.16$; $F(2, 324) = 0.17$, $p =$

543 0.844]. We inspected RT differences between approach and avoid trials and there was *strong*

544 evidence for slower RTs on pull compared to push actions $[BF_{10} = 11.32$; $t(162) = 2.95$, $p =$

545 0.004, $d = 0.23$, 95% CI for $d = 0.08$, 0.39]. Together these results may indicate fatigue

546 effects associated with bio-mechanical costs (e.g., arm flexion muscle group activation).

Reliability of calculated bias scores.   Given the absence of evidence for baseline approach-avoidance bias scores deviating from zero and the variability in the distribution of score from negative to positive bias (see Figure 2), we checked the test-retest reliability of the task measures. The present study included a training intervention between time points, but since there was no effect of training on AAT bias scores results would still be informative. Considering that the the interval between pre- and post-training AAT blocks was very short we considered that a test-retest reliability (or stability) coefficient r greater than 0.7 would be adequate in this context. Test-retest reliability was therefore assessed via correlation pairs for AAT bias scores at baseline (AAT_{pre}) and after training (AAT_{post}), as shown in Figure **??** (panel C). Consistent with our previous analyses of baseline bias scores (see *Baseline approach bias scores*), test-retest reliability was examined for both completion time and initiation time AAT bias scores. Bayesian correlation pairs with the default prior (stretched beta with $\gamma=1$; Wagenmakers, Verhagen, and Ly (2016)) were used for these analyses. As expected, we found *extreme* evidence for a positive linear relationship between completion time bias scores for pre- and post- training blocks, but the correlation coefficient (Pearson's rho) was only 0.41, indicating that there test-retest reliability of AAT bias scores based on completion times was poor [$\log(BF_{10}) = 12.95$, $p < .001$, 95% CI for $r = 0.30$, 1]. As discussed earlier, bias scores based on median completion times could be affected by noise in motor times and scores based on median initiation times would better reflect underlying cognitive processes, such as *automatic* action tendencies. The test-retest reliability for bias scores based on initiation times however was only slightly better compared to completion time bias scores with a stability coefficient of 0.53 [$\log(BF_{10}) = 24.34$, $p < .001$, 95% CI for $r = 0.43$, 1].

Devaluation trends across training conditions.   As explained in Figure 3, there is a general trend of devaluation in the data for all training conditions from pre- to post-training. We tested for these differences directly and obtained conclusive evidence that within each training condition cell, there is a negative change in mean liking ratings from

574 pre- to post-training, as observed in Figure 3. The control (filler) foods should be unaffected

575 in terms of affective evaluation changes, but participants rated filler foods more negatively

576 after training ($M = 68.55$, $SD = 15.81$) relative to baseline ($M = 71.16$, $SD = 14.80$) [BF_{10}

577 $= 156.54$, $W = 8436.00$, $p < .001$, $r = 0.29$, 95% CI for $r = 0.12$, 0.45]. Contrary to

578 predictions about the increase in positive evaluations for go foods (relative to control),

579 within that condition cell the evaluations of go foods were less positive after training ($M =$

580 67.42, $SD = 16.85$) compared to before ($M = 70.29$, $SD = 16.80$) [$BF_{10} = 84.52$, $W =$

581 7898.50, $p < .001$, $r = 0.31$, 95% CI for $r = 0.13$, 0.46]. The effect was greater for no-go

582 foods, but this is the only data trend that was theoretically consistent with effects of

583 training. Participants provided less positive ratings for no-go foods after training ($M =$

584 68.83, $SD = 16.81$) compared to before ($M = 72.99$, $SD = 15.38$) [$BF_{10} = 211398.68$, $W =$

585 9947.50, $p < .001$, $r = 0.49$, 95% CI for $r = 0.34$, 0.61].

586 **General linear model of food choice data.** As shown in Figure 4 and as

587 expected for counts data, food choice probabilities were not normally distributed and the

588 inferences based on paired t-tests would need to be validated further. Choice count data

589 from the impulsive food choice task were modelled using a general linear model (GLM) in R

590 (R Core Team, 2017). The error term of the model was specified with a Poisson distribution

591 and the link function log-transformed the linear predictor within the model (i.e., logarithms

592 of fitted means). The only predictor in this model was the training condition (i.e.., go, no-go,

593 or filler foods), with filler, or control, foods set as the reference point (i.e., intercept).

594 Diagnostic plots showed mild violations of the assumptions of homoskedasticity and

595 normality of residuals and thus robust standard errors for the parameter estimates were

596 computed (Cameron & Trivedi, 2009). An overdispersion test (Cameron & Trivedi, 2005;

597 Kleiber & Zeileis, 2008) showed that true dispersion was not greater than 1 and the

598 goodness-of-fit chi-squared test was not statistically significant, indicating that the model

599 had a good fit (Residual variance $= 436.15$, df $= 486$, $p = 0.949$). The GLM results are

600 consistent with the pre-registered statistical test results (see Table 1). The model showed

that impulsive choice probability for no-go foods was 0.53 times the choice probability for control foods [Estimate (log) = -0.617, Robust SE = 0.119, $p < .001$, 95% CI = -0.851, -0.383)]. The probability of choosing a go food after training was 1.22 times the probability of choosing a control food [Esimate (log) = 0.197, Robust SE = 0.091, $p = 0.030$, 95% CI = 0.019, 0.374].

Robust statistics. For certain pre-registered paired comparisons, where the difference scores were found to violate the normality assumption (Shapiro-Wilk test with p leq .005), the possibility of biased effect size esimates was considered and therefore robust statistics are also reported (Lakens, 2015). H2a and H2b have been ommitted as robust analyses have already been implemented above (see *General linear model of food choice data*). We used Yuen's method of comparing trimmed means via the WRS2 package, with the recommended percentage of 20% trimming from both tails of the distribution (Mair & Wilcox, 2019; Wilcox & Tian, 2011; Yuen, 1974). The exploratory measure of effect size, represented by ξ is provided and can be conventionally interpreted as small, median and large at 0.15, 0.35 and 0.50 (Mair & Wilcox, 2019). The null hypothesis in Yuen's test for paired sample comparisons is that there is no difference between the trimmed means ($\mu_{t1} = \mu_{t2}$). The test results are shown in Table 3 and are consistent withpre-registered findings. The explanatory effect sizes do not deviated in their interpretation from Cohen's d values presented in Table **??** for frequentists paired-samples t-tests, as the observed effects are small for both H3a and H4a.

Discussion

No effects of training on automatic action tendencies

The primary aim of the study was to investigate whether inhibitory control training (ICT) can have an indirect effect on automatic action tendencies. It was hypothesized that

Table 3

Yuen's tests of trimmed mean differences for paired comparisons that violated the normality assumption

	$t(98)$	p	MD_t	95% CI Lower	Upper	ξ	Comparison to confirmatory test	Effect size interpretation
H1a	-0.90	0.369	-5.12	-16.37	6.14	0.07	Consistent	None
H3a	-1.67	0.09818	-1.03	-2.25	0.19	0.10	Consistent	Very small
H3b	-0.09	0.9319	-0.06	-1.50	1.38	0.01	Consistent	None
H4a	-3.11	0.002419	-0.01	-0.01	0.00	0.20	Consistent	Small

Note. The degrees of freedom for the robust t statistic are 98 because of trimming at 20% for both tails of the distribution (i.e., N=99). Comparison to confirmatory test refers to whether or not the results from Yuen's tests are consistent with the pre-registered, confirmatory test results. MD_t: trimmed mean difference; ξ: explanatory measure of effect size

approach bias for unhealthy foods associated with a no-go response during go/no-go training (i.e., response inhibition) would be reduced compared to foods that were paired with both go and no-go responses with an equal probability (control foods). Automatic action tendencies were indirectly measured using a variant of the approach-avoidance task (AAT) that includes a "zooming" feature for push/pull actions (Neumann & Strack, 2000) of the computer mouse and requires participants to judge the orientation of the presented picture (C. E. Wiers et al., 2013). We calculated approach-avoidance bias scores from AAT blocks before and after training by subtracting median RTs on approach trials (pull action) from median RTs on avoid trials (push action). Positive scores would indicate an approach bias towards unhealthy foods. For our primary outcome measures, we compared the change in bias scores from pre-to post-training across training conditions. Our results show that ICT did not have an effect on automatic action tendencies, as there was *moderate* evidence that approach bias for no-go foods was not reduced relative to control foods after training (H1a) and *strong* evidence that approach bias for go foods was not increased compared to control foods after training (H1b). To our knowledge, the AAT has not yet been employed as an outcome

measure in ICT studies utilising the go/no-go training paradigm and the null findings have shed light into methodological and theoretical issues to be explored further, which will be discussed after all study outcomes have been thoroughly described.

Response inhibition & impulsive food choices

As a secondary outcome of training, we tested whether go/no-go training had an effect on impulsive food choices. In an adapted food choice task (Veling et al., 2013a) participants had ten seconds to choose three food stimuli from all training conditions were and evaluations of unhealthy foods stimuli (i.e., liking). Pre-registered analyses showed that the probability of choosing a no-go food in an impulsive food choice task was reduced compared to the probability of choosing a control food (H2a). Meanwhile, there was no difference between the probability of choosing a go food relative to the probability of choosing a control food (H2b). The findings that ICT can have an effect on impulsive food choices is consistent with previous studies that have used both go/no-go and stop-signal task paradigms (Veling et al., 2013b, 2013a), but cannot be directly compared to experiments involving cue-approach training, which involves responding to go items in response to a cue or signal. These studies have found increased food choices for go food items (Schonberg et al., 2014; Veling, Chen, et al., 2017), but the cued-approach different attentional focus in this paradigm, oriented towards go foods, we cannot state that the lack of increased choice probability for go foods in this study contradicts previous findings.

Devaluation effects & design limitations

Another important training outcome which was also treated as manipulation check for our ICT paradigm was the change in food evaluations from pre-to post-training. According to the BSI theory of stimulus devaluation, as already discussed, successful inhibition of

responses on signal trials is facilitated by an underlying devaluation process for appetitive foods, whereby approach bias for these foods is reduced (Veling et al., 2008, 2017). Consistent with previous findings where go/no-go training led to robust food devaluation effects (see Chen et al., 2016 for a series of pre-registered experiments), we expected that the change in mean tastiness ratings (i.e., food liking) for no-go foods from pre-to post-training would be reduced compared to the change in ratings for filler foods. Pre-registered analyses showed only *anecdotal* evidence that no-go foods were rated less positively after training compared to filler foods (H3a). Similarly, participants did not show increased liking for go foods relative to the filler foods, from pre-to post-training (H3b). It should be noted that filler foods which were associated with both go and no-go responses (50:50) are not an ideal control for devaluation effects, as Chen et al. (2016) correctly point out that two contrasts should be tested for proper inferences to be made. They compared changes in evaluation for no-go foods compared to changes for go foods as well as changes for untrained food stimuli, which were never included in training. All food stimuli sets were matched in valence at the beginning of the study. Contrary to that design, we only presented a fixed set of unhealthy foods which were considered appetitive (e.g., pizza, cake, crisps) and this is a viable limitation of the study with regards to the examination of devaluation effects.

Exploratory analyses further showed that while on average no-go foods were rated less positively after training compared to before, there was a general devaluation trend for both go and control foods (see *Devaluation trends across training conditions*). We therefore considered the possibility that over-exposure to food stimuli from all training conditions during the phases of the study (pre-training, training, and post-training) could have had a viable habituation effect on participants' affective evaluations of any presented foods at the end of the study. For this reason and the robustness of the design employed by Chen et al. (2016) discussed above, we recommend that future studies only present outcome measures, such as the AAT only after training for the first time, which can reduce habituation effects, and that untrained food stimuli are included as additional controls. The recommended

690 design can also increase the number of observations without the need for great data

691 reduction, such as calculating a difference score of the AAT_{pre} and AAT_{post} bias scores which

692 are also represented by difference scores between median RTs in approach and avoid trials.

693 The absence of devaluation effects specific to no-go foods could not be attributed to

694 ineffective training, as our second manipulation check for contingency learning during the

695 go/no-go task was positive. We found *extreme* evidence that GoRTs on correct no-signal

696 trials were reduced for go foods compared to control foods and that the percentage of correct

697 responses on signal trials were greater for no-go foods relative to control foods.

698 **Methodological considerations for the approach-avoidance task**

699   First, we showed that overall baseline bias scores did not deviate from zero (see

700 *Baseline approach bias scores*) which suggests that either participants did not have any

701 approach bias for the selected foods or the employed variant of the AAT was not sensitive

702 enough to capture both baseline bias and potential indirect effects of training. Sub-group

703 analyses showed that even when participants had positive bias scores for unhealthy foods,

704 which reflect existing approach bias, there were still no effects of training on automatic

705 action tendencies (see *Sub-group analysis*). We then examined the test-retest reliability of

706 the calculated bias scores for pre-and post-training AAT blocks and the stability coefficients

707 showed poor reliability (see *Reliability of calculated bias scores*). Test-retest reliability was

708 slightly improved when bias scores were calculated using initiation times, instead of

709 completion times, which may be due to individual differences in action-associated motor

710 demands (i.e.,biceps and triceps muscle activation for pushing/pulling the computer mouse).

711 We suggest that initiation time would be a more reliable measure to use for calculation AAT

712 bias scores (e.g., see Seibt, Häfner, & Deutsch, 2007). It is also unclear whether any time of

713 response latency measure derived from sensorimotor tasks, such as the AAT variant

714 employed in this study, can be indicative of approach-avoidance bias, since the role of arm

movements in these motivational processes has recently been questioned for the controversial replicability of findings and the importance of whole-body movements in real-world approach-avoidance behaviours (Rougier et al., 2018). An element of the AAT variant that could have affected the reliability of the bias scores was the use of a computer mouse instead of a joystick, as in the seminal version of this paradigm (Rinck & Becker, 2007; Wiers et al., 2009). However, any motor demand differences between the joystick and computer mouse would not affect the initiation times, which were also found to have questionable test-retest reliability. We should highlight the importance of variability in methodology for the measurement of motivational bias and that different parameters would also need to be examined further when the AAT is applied in the food domain (e.g., explicit vs implicit task instructions; see Phaf et al., 2014 for meta-analysis).

Certainly, there are methodological issues with the application of the AAT as an indirect measure of approach bias in ICT studies that have strict comparisons for outcomes based only food stimuli. For example, one approach for making inferences based on AAT scores is to calculate a *relative* bias based on the push-pull median RT differences from one category (e.g., alcohol) versus another (e.g., non-alcohol), as reported in previous literature (Sharbanee, Stritzke, Jamalludin, & Wiers, 2014; van Deursen, Salemink, Smit, Kramer, & Wiers, 2013). However, in the present study design the AAT included only sweet and savoury unhealthy foods and bias scores were based on the differences between push and pull actions. The control for relative bias scores should be specific to the question of interest, such as healthy foods if approach bias and healthiness is to be examined or non-food stimuli if a general food approach bias is to be inferred. In any case, the control stimuli need to be matched in liking (i.e., hedonic value) so that any resulting bias scores reflect differences in motivational bias and not affective/hedonic bias (Kemps & Tiggemann, 2015).

mention that we found evidence for fatigue effects associated with pull actions after training- triceps

Concluding remarks & future directions

From a theoretical standpoint, we assumed that there could be a link between stimulus devaluation during ICT training and automatic action tendencies. If a tendency to approach an appetitive food is reduced during go/no-go training in order for response inhibition to be successful, the approach bias towards food stimuli associated with signal trials could be indirectly affected by this process. The absence of devaluation effects in this study could therefore explain the finding that ICT did not have any influence on automatic action tendencies. Nevertheless, there are several methodological issues regarding the application of the AAT as an indirect measure of motivational bias that need to be considered before drawing any conclusions and would need to be addressed in future studies.
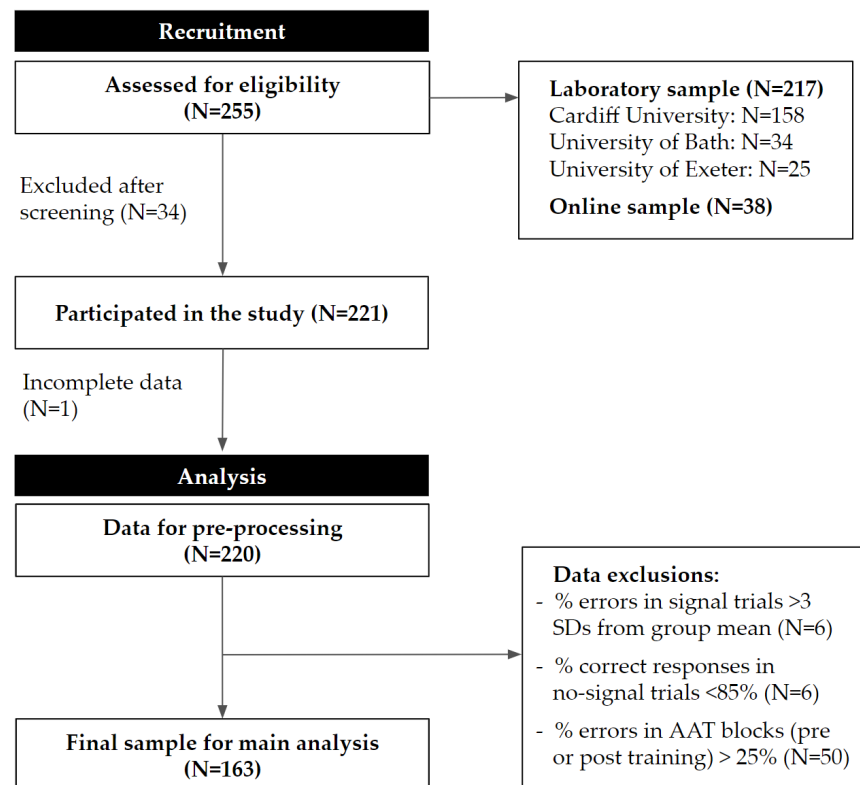
Appendix A

Recruitment & data exclusions

Figure A1. **Flow diagram of recruitment and participant-level data exclusions.** There were 255 individuals recruited and assessed for eligibility across laboratory sites and online via personal communication. 34 participants were excluded after screening for not meeting the advertised inclusion/exclusion criteria and we obtained datasets from 221 participants. The online sample was recruited by the University of Bath and University of Exeter. One participant was excluded for providing incomplete data and 220 datasets were submitted for pre-processing and inspection. There we no participants with a mean reaction time on no-sginal trials (GoRT) greater than three standard deviations (SDs) from the group mean and there we no cases of consistently missed (i.e., default option of 50) responses on food rating trials. Six participants had a percentage of errors in signal trials was greater than three SDs from the group mean and six participants also had a percentage of correct responses in no-signal trials lower than 85%. Please note that some participants met more than one exclusion criterion. 50 participants were excluded as their percentage of errors in either the pre- or post-training approach-avoidance task (AAT) blocks was greater than 25%. The final sample consisted of 163 participants.

Appendix B

Sample characteristics & questionnaire measures

751 All sample characteristics, apart from gender and hours since last meal, are presented in the

752 Table B2 together with total scores from relevant questionnaire measures, as described in the

753 *Survey & Questionnaires* section. Descriptive statistics of the questionnaire scores can be

754 found in Table B1. Pearson's *r* coefficients were conventionally interpreted as small, medium

755 and large at 0.10, 0.30 and 0.50. As it would be expected for the Food Cravings

756 Questionnaire Trait- reduced (FCQ-T-r) measure, there was a positive correlation, although

757 small, with BMI as well as medium-to-large positive correlations with total scores on the

758 Barratt Impulsivity Scale (BIS) and Perceived Stress Scale (PSS). Trait food cravings

759 negatively correlated with stop control, as measured by the Stop Control Scale (SCS) and

760 the food subscale of the Delaying Gratification Inventory (DGI).

Table B1

Descriptive statistics of questionnaire scores from the final sample

	FCQ-T-r total	BIS total	PSS total	SCS total	DGI - food
Mean	45.362	32.773	19.896	40.951	22.245
Median	45.000	32.000	19.000	41.000	22.000
Standard Deviation	9.997	5.751	6.298	7.665	4.677
Minimum	20.000	21.000	4.000	20.000	10.000
Maximum	70.000	51.000	38.000	57.000	34.000

Note. For descriptions and abbreviations, please see the *Survey & Questionnaires* section.

Table B2

Correlation matrix for sample characteristics and questionnaire measures

		1.	2.	3.	4.	5.	6.	7.	8.
1. Age	r	–							
	$\log(BF_{10})$	–							
	p	–							
2. Hunger	r	-0.064	–						
	$\log(BF_{10})$	0.352	-2.322	–					
	p	0.420	–						
3. BMI	r	0.182	0.001	–					
	$\log(BF_{10})$	0.352	-2.322	–					
	p	0.020	0.987	–					
4. FCQ-T-r	r	-0.098	0.203	0.246*	–				
	$\log(BF_{10})$	-1.554	1.034	2.655	–				
	p	0.213	0.009	0.002	–				
5. BIS	r	-0.089	0.129	0.161	0.491*	–			
	$\log(BF_{10})$	-1.690	-0.991	-0.245	19.572	–			
	p	0.259	0.101	0.041	< .001	–			
6. PSS	r	-0.138	0.040	0.125	0.462*	0.316*	–		
	$\log(BF_{10})$	-0.782	-2.195	-1.067	16.754	6.043	–		
	p	0.078	0.611	0.111	< .001	< .001	–		
7. SCS	r	0.176	-0.085	-0.122	-0.374*	-0.721*	-0.260*	–	
	$\log(BF_{10})$	0.181	-1.747	-1.133	9.630	56.042	3.247	–	
	p	0.025	0.281	0.121	< .001	< .001	< .001	–	
8. DGI-food	r	0.075	-0.161	-0.189	-0.612*	-0.433*	-0.226	0.376*	–
	$\log(BF_{10})$	-1.878	-0.237	0.577	35.070	14.168	1.849	9.777	–
	p	0.343	0.040	0.016	< .001	< .001	0.004	< .001	–

Note. Age was self-reported in years and hunger ratings ranged from 1="Not at all" to 9="Very". Body-mass index (BMI; kg/m^2); * Supported correlations at $BF_{10} > 10$

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. https://doi.org/10.1016/j.appet.2016.11.014

Allen, M., Poggiali, D., Whitaker, K., Marshall, T., & Kievit, R. (2019). Raincloud plots: A multi-platform tool for robust data visualization [version 1; peer review: 2 approved]. *Wellcome Open Research*, *4* (63). https://doi.org/10.12688/wellcomeopenres.15191.1

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. (2018). RainCloudPlots tutorials and codebase. Zenodo. https://doi.org/10.5281/zenodo.1402959

Allom, V., Mullan, B., & Hagger, M. (2016). Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychol. Rev.*, *10* (2), 168–186. https://doi.org/10.1080/17437199.2015.1051078

Bartholdy, S., Dalton, B., O'Daly, O. G., Campbell, I. C., & Schmidt, U. (2016). A systematic review of the relationship between eating, weight and inhibitory control using the stop signal task. *Neurosci. Biobehav. Rev.*, *64*, 35–62. https://doi.org/10.1016/j.neubiorev.2016.02.010

Becker, D., Jostmann, N. B., Wiers, R. W., & Holland, R. W. (2015). Approach avoidance training in the eating domain: Testing the effectiveness across three single session studies. *Appetite*, *85* (June 2015), 58–65. https://doi.org/10.1016/j.appet.2014.11.017

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., . . . Johnson, V. E. (2017). Redefine statistical significance. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-017-0189-z

Bradley, B., Field, M., Healy, H., & Mogg, K. (2008). Do the affective properties of
        smoking-related cues influence attentional and approach biases in cigarette smokers?
        Journal of Psychopharmacology, *22*(7), 737–745.
        https://doi.org/10.1177/0269881107083844

Brignell, C., Griffiths, T., Bradley, B. P., & Mogg, K. (2009). Attentional and approach
        biases for pictorial food cues. Influence of external eating. *Appetite*, *52*(2), 299–306.
        https://doi.org/10.1016/j.appet.2008.10.007

Brockmeyer, T., Hahn, C., Reetz, C., Schmidt, U., & Friederich, H. C. (2015). Approach
        Bias Modification in Food Craving - A Proof-of-Concept Study. *Eur. Eat. Disord.
        Rev.*, *23*(5), 352–360. https://doi.org/10.1002/erv.2382

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., &
        Munafò, M. R. (2013). Power failure: Why small sample size undermines the
        reliability of neuroscience. *Nat. Rev. Neurosci.*, *14*(5), 365–376.
        https://doi.org/10.1038/nrn3475

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications.*
        Cambridge, New York: Cambridge University Press.

Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics Using Stata.* College Station,
        TX: Stata Press.

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*(12), 1687–1701.
        https://doi.org/10.1037/xge0000236

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress.
        Journal of Health and Social Behavior, *24*(4), 385–396.

De Boer, B. J., van Hooft, E. A. J., & Bakker, A. B. (2011). Stop and start control: A distinction within self-control. *European Journal of Personality*, *25*(5), 349–362. https://doi.org/10.1002/per.796

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Field, M., Eastwood, B., Bradley, B. P., & Mogg, K. (2006). Selective processing of cannabis cues in regular cannabis users. *Drug and Alcohol Dependence*, *85*(1), 75–82. https://doi.org/10.1016/j.drugalcdep.2006.03.018

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*(1), 66–73. https://doi.org/10.1016/j.appet.2006.11.008

Hall, P. A. (2012). Executive control resources and frequency of fatty food consumption: Findings from an age-stratified community sample. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *31*(2), 235–241. https://doi.org/10.1037/a0025407

Havermans, R. C. (2013). Pavlovian Craving and Overeating: A Conditioned Incentive Model. *Current Obesity Reports*, *2*(2), 165–170. https://doi.org/10.1007/s13679-013-0053-z

Hoerger, M., Quirk, S. W., & Weed, N. C. (2011). Development and validation of the Delaying Gratification Inventory. *Psychological Assessment*, *23*(3), 725–738. https://doi.org/10.1037/a0023286

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

temptations. *Appetite*, *56*(2), 345–349. https://doi.org/10.1016/j.appet.2010.12.017

Houben, K., & Jansen, A. (2015). Chocolate equals stop: Chocolate-specific inhibition training reduces chocolate intake and go associations with chocolate. *Appetite*, *87*, 318–323. https://doi.org/10.1016/j.appet.2015.01.005

Houben, K., Nederkoorn, C., & Jansen, A. (2012). Too tempting to resist? Past success at weight control rather than dietary restraint determines exposure-induced disinhibited eating. *Appetite*, *59*(2), 550–555. https://doi.org/10.1016/j.appet.2012.07.004

Jasinska, A. J., Yasuda, M., Burant, C. F., Gregor, N., Khatri, S., Sweet, M., & Falk, E. B. (2012). Impulsivity and inhibitory control deficits are associated with unhealthy eating in young adults. *Appetite*, *59*(3), 738–747. https://doi.org/10.1016/j.appet.2012.08.001

JASP Team. (2018). JASP (Version 0.10.0)[Computer software].

Jones, A., Di Lemma, L. C., Robinson, E., Christiansen, P., Nolan, S., Tudur-Smith, C., & Field, M. (2016). Inhibitory control training for appetitive behaviour change: A meta-analytic investigation of mechanisms of action and moderators of effectiveness. *Appetite*, *97*, 16–28. https://doi.org/10.1016/j.appet.2015.11.013

Jones, A., Hardman, C. A., Lawrence, N., & Field, M. (2018). Cognitive training as a potential treatment for overweight and obesity: A critical review of the evidence. *Appetite*, *124*, 50–67. https://doi.org/10.1016/j.appet.2017.05.032

Kakoschke, N., Kemps, E., & Tiggemann, M. (2017a). Approach bias modification training and consumption: A review of the literature. *Addictive Behaviors*, *64*, 21–28. https://doi.org/10.1016/j.addbeh.2016.08.007

Kakoschke, N., Kemps, E., & Tiggemann, M. (2017b). The effect of combined avoidance and

control training on implicit food evaluation and choice. *J. Behav. Ther. Exp.*
Psychiatry, *55*, 99–105. https://doi.org/10.1016/j.jbtep.2017.01.002

Kakoschke, N., Kemps, E., Tiggemann, M., Kakoschke, N., Kemps, E., & Tiggemann, M.
(2015). Combined effects of cognitive bias for food cues and poor inhibitory control
on unhealthy food intake. *Appetite*, *87*(JANUARY), 358–364.
https://doi.org/10.1016/j.appet.2015.01.004

Kemps, E., & Tiggemann, M. (2015). Approach bias for food cues in obese individuals.
Psychol. Heal., *30*(3), 370–380. https://doi.org/10.1080/08870446.2014.974605

Kemps, E., Tiggemann, M., Martin, R., & Elliott, M. (2013). Implicit approachAvoidance
associations for craved food cues. *J. Exp. Psychol. Appl.*, *19*(1), 30–38.
https://doi.org/10.1037/a0031626

Kleiber, C., & Zeileis, A. (2008). *Applied Econometrics with R*. New York: Springer-Verlag.

Lakens, D. (2015). The 20% Statistician: The perfect t-test. *The 20% Statistician.*

Lawrence, N. S., Hinton, E. C., Parkinson, J. A., & Lawrence, A. D. (2012). Nucleus
accumbens response to food cues predicts subsequent snack consumption in women
and increased body mass index in those with reduced self-control. *NeuroImage*, *63*(1),
415–422. https://doi.org/10.1016/j.neuroimage.2012.06.070

Lawrence, N. S., O'Sullivan, J., Parslow, D., Javaid, M., Adams, R. C., Chambers, C. D., . . .
Verbruggen, F. (2015). Training response inhibition to food is associated with weight
loss and reduced energy intake. *Appetite*, *95*, 17–28.
https://doi.org/10.1016/j.appet.2015.06.009

Lawrence, N. S., Verbruggen, F., Morrison, S., Adams, R. C., & Chambers, C. D. (2015).
Stopping to food can reduce intake. Effects of stimulus-specificity and individual

differences in dietary restraint. *Appetite*, *85*, 91–103.

https://doi.org/10.1016/j.appet.2014.11.006

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course.*

Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Mair, P., & Wilcox, R. (2019). Robust statistical methods in R using the WRS2 package.

Behavior Research Methods. https://doi.org/10.3758/s13428-019-01246-w

Meule, A., Hermann, T., & Kübler, A. (2014). A short version of the food cravings

questionnaire-trait: The FCQ-T-reduced. *Front. Psychol.*, *5*(MAR), 1–10.

https://doi.org/10.3389/fpsyg.2014.00190

Nederkoorn, C., Coelho, J. S., Guerrieri, R., Houben, K., & Jansen, A. (2012). Specificity of

the failure to inhibit responses in overweight children. *Appetite*, *59*(2), 409–413.

https://doi.org/10.1016/j.appet.2012.05.028

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 : Official Journal of the Division of Health Psychology, American

Psychological Association*, *29*, 389–393. https://doi.org/10.1037/a0019921

Neumann, R., & Strack, F. (2000). Approach and Avoidance: The Influence of

Proprioceptive and Exteroceptive Cues on Encoding of Affective Information. *J.

Personal. Soc. Psychol.*, *79*(1), 39–48. https://doi.org/10.1037//0022-3514.79.1.39

Phaf, R. H., Mohr, S. E., Rotteveel, M., & Wicherts, J. M. (2014). Approach, avoidance,

and affect: A meta-analysis of approach-avoidance tendencies in manual reaction time

tasks. *Front. Psychol.*, *5*(378), 1–16. https://doi.org/10.3389/fpsyg.2014.00378

R Core Team. (2017). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Richard D. Morey. (2015). On verbal categories for the interpretation of Bayes factors.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 105–120. https://doi.org/10.1016/j.jbtep.2006.10.001

Rouder, J. N., Engelhardt, C. R., McCabe, S., & Morey, R. D. (2016). Model comparison in ANOVA. *Psychon. Bull. Rev., 23*(6), 1779–1786. https://doi.org/10.3758/s13423-016-1026-5

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Rougier, M., Muller, D., Ric, F., Alexopoulos, T., Batailler, C., Smeding, A., & Aubé, B. (2018). A new look at sensorimotor aspects in approach/avoidance tendencies: The role of visual whole-body movement information. *Journal of Experimental Social Psychology, 76*, 42–53. https://doi.org/10.1016/j.jesp.2017.12.004

RStudio Team. (2016). *RStudio: Integrated Development Environment for R.* Boston, MA: RStudio, Inc.

Schonberg, T., Bakkour, A., Hover, A. M., Mumford, J. A., Nagar, L., Perez, J., & Poldrack, R. A. (2014). Changing value through cued approach: An automatic mechanism of behavior change. *Nat. Neurosci., 17*(4), 625–630. https://doi.org/10.1038/nn.3673

Schumacher, S. E., Kemps, E., & Tiggemann, M. (2016). Bias modification training can alter

approach bias and chocolate consumption. *Appetite*, *96*, 219–224.

https://doi.org/10.1016/j.appet.2015.09.014

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*(2), 359–379. https://doi.org/10.1002/ejsp.365

Sharbanee, J. M., Stritzke, W. G., Jamalludin, M. E., & Wiers, R. W. (2014).

Approach-alcohol action tendencies can be inhibited by cognitive load.

Psychopharmacology (Berl)., *231*(5), 967–975.

https://doi.org/10.1007/s00213-013-3318-z

Spinella, M. (2007). Normative Data and a Short Form of the Barratt Impulsiveness Scale.

International Journal of Neuroscience, *117*(3), 359–368.

https://doi.org/10.1080/00207450600588881

Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior.

Personality and Social Psychology Review, *8*(3), 28.

van Deursen, D. S., Salemink, E., Smit, F., Kramer, J., & Wiers, R. W. (2013). Web-based

cognitive bias modification for problem drinkers: Protocol of a randomised controlled

trial with a 2x2x2 factorial design. *BMC Public Health*, *13*, 674.

https://doi.org/10.1186/1471-2458-13-674

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

approach tendencies towards food. *Appetite*, *55*(1), 30–36.

https://doi.org/10.1016/j.appet.2010.03.007

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

responses toward palatable foods. *Behav. Res. Ther.*, *49*(11), 771–780.

https://doi.org/10.1016/j.brat.2011.08.005

Veling, H., Aarts, H., & Stroebe, W. (2013a). Stop signals decrease choices for palatable foods through decreased food evaluation. *Front. Psychol.*, *4*(875), 1–7. https://doi.org/10.3389/fpsyg.2013.00875

Veling, H., Aarts, H., & Stroebe, W. (2013b). Using stop signals to reduce impulsive choices for palatable unhealthy foods. *Br. J. Health Psychol.*, *18*(2), 354–368. https://doi.org/10.1111/j.2044-8287.2012.02092.x

Veling, H., Chen, Z., Tombrock, M. C., M. Verpaalen, I. a., Schmitz, L. I., Dijksterhuis, A., & Holland, R. W. (2017). Training Impulsive Choices for Healthy and Sustainable Food. *J. Exp. Psychol. Appl.*, *23*(1), 1–14. https://doi.org/10.1037/xap0000112

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*(4), 1013–1019. https://doi.org/10.1016/j.jesp.2008.03.004

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. *Curr. Addict. Reports*, *4*(1), 35–41. https://doi.org/10.1007/s40429-017-0131-5

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. https://doi.org/10.1016/j.appet.2014.03.014

Verbruggen, F., Best, M., Bowditch, W. A., Stevens, T., & McLaren, I. P. (2014). The inhibitory control reflex. *Neuropsychologia*, *65*(312445), 263–278. https://doi.org/10.1016/j.neuropsychologia.2014.08.014

Wagenmakers, E.-J., Verhagen, J., & Ly, A. (2016). How to quantify the evidence for the

absence of a correlation. *Behavior Research Methods*, *48*(2), 413–426.
https://doi.org/10.3758/s13428-015-0593-0

WHO. (2018). Obesity and overweight. *World Health Organization.*
https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight.

Wiers, C. E., Kühn, S., Javadi, A. H., Korucuoglu, O., Wiers, R. W., Walter, H., . . .
Bermpohl, F. (2013). Automatic approach bias towards smoking cues is present in
smokers but not in ex-smokers. *Psychopharmacology*, *229*(1), 187–197.
https://doi.org/10.1007/s00213-013-3098-5

Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining
Automatic Action Tendencies Changes Alcoholic Patients' Approach Bias for Alcohol
and Improves Treatment Outcome. *Psychol. Sci.*, *22*(4), 490–497.
https://doi.org/10.1177/0956797611400615

Wiers, R. W., Gladwin, T. E., & Rinck, M. (2013). Should we train alcohol-dependent
patients to avoid alcohol? *Front. Psychiatry*, *4*(May), 33.
https://doi.org/10.3389/fpsyt.2013.00033

Wiers, R. W., Rinck, M., Dictus, M., & Van Den Wildenberg, E. (2009). Relatively strong
automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele.
Genes, Brain Behav., *8*(1), 101–106.
https://doi.org/10.1111/j.1601-183X.2008.00454.x

Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic
action-tendencies to approach alcohol in hazardous drinkers. *Addiction*, *105*(2),
279–287. https://doi.org/10.1111/j.1360-0443.2009.02775.x

Wilcox, R. R., & Tian, T. S. (2011). Measuring effect size: A robust heteroscedastic
approach for two or more groups. *Journal of Applied Statistics*, *38*(7), 1359–1368.

996   https://doi.org/10.1080/02664763.2010.498507

997   Wittekind, C. E., Feist, A., Schneider, B. C., Moritz, S., & Fritzsche, A. (2015). The

998   approach-avoidance task as an online intervention in cigarette smoking: A pilot study.

999   *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 115–120.

1000   https://doi.org/10.1016/j.jbtep.2014.08.006

1001   Yuen, K. K. (1974). The two-sample trimmed t for unequal population variances.

1002   *Biometrika, 61*(1), 165–170. https://doi.org/10.1093/biomet/61.1.165

1003   −>