No effects of explicit approach-avoidance training on immediate consumption of soft drinks

Anand Krishna & Andreas B. Eder

University of Würzburg, Germany

In press, Appetite

Author Note:

Correspondence concerning this article should be addressed to AE, Department of Psychology, University of Würzburg, Röntgenring 10, 97070 Würzburg, Germany.

Electronic mail may be sent via Internet to andreas.eder@uni-wuerzburg.de.

Phone: +49 931 31 83336, Fax: +49 931 31 82812.

This research was supported by grant ED 201/3-1 of the German Research Foundation (DFG) to AE. The DFG had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

Abstract

Four experiments (n = 300) examined motivational effects of approach-avoiding training (AAT) procedures on consumption of sugary soft drinks, implicit preferences and explicit preferences. Experiments varied in the number of training trials, the implementation of approach-avoidance goals during the training, and the frequency and timing of the consumption measure. AAT had no effects on any measure, and Bayesian analyses provided substantial evidence for a null model of AAT effects. A manipulation check showed that AAT affected behavioral tendencies towards the drinks in line with the training procedure (Experiment 3). It is concluded that explicit training of approach and avoidance reactions to soft drinks is not an effective procedure to modify immediate consumption of that drinks. Possible reasons and differences to previous AAT studies are discussed.

Keywords: soft drink consumption; approach-avoidance training; preference;

Obesity and related health problems such as diabetes are widely acknowledged as a prevalent and rising health risk in Western (Ogden, Carroll, Kit, & Flegal, 2014). Although many factors contribute to this rising tide, increasing consumption of sweetened soft drinks has been identified as an important contributor to negative health outcomes (Mensink et al., 2018; Vartanian, Schwartz, & Brownell, 2007). Even though the health risks are established, campaigns to change people's attitudes and behavior towards soft drinks via education and thereby mitigate some of their negative effects have shown limited success (Block, Chandra, McManus, & Willett, 2010). Therefore, other, more effective motivational interventions are needed to address this issue.

In recent years, psychologists have invented new computerized interventions that aim to change automatic or implicit motivational processes involved in consumptive behaviors (Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013). These new interventions are here collectively referred to as *approach-avoidance training* (AAT), because they seek to modify motivational action tendencies by a retraining of approach- and avoidance-related behavioral responses. The rationale of AAT is that approach-avoidance tendencies can be changed with the repeated execution of a behavior that is congruent or incongruent with a motivational tendency to approach and avoid. The training procedure has its precursor in attention bias modification (ABM), which modifies covert attentional processes by directing attention repeatedly towards and/or away from specific stimuli (Hakamata et al., 2010; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002). In contrast to ABM, however, AAT aims to change overt approach and avoidance responses and the motivational processes that energize these responses.

Most AAT studies have used movements of a joystick lever for a retraining of approach-avoidance tendencies. Most procedures assert that pulling a lever towards the body is associated with an approach motivation, while pushing a lever away from oneself is associated with avoidance (Eder & Rothermund, 2008). Executing a lever push or pull in response to a specific stimulus activates the associated motivational orientation, and through this link, a motivation to approach or avoid the stimulus. After sufficient training, the stimulus becomes associated with the motivational orientation that was activated by the trained response.

Consistent with this theorizing, many studies obtained evidence that a training of approach- and avoidance-related action tendencies can affect social (e.g., Kawakami, Phills, Steele, & Dovidio, 2007), emotional (e.g., Amir, Kuckertz, & Najmi, 2013), and consumptive outcomes (e.g., Schumacher, Kemps, & Tiggemann, 2016). For instance, in one study participants sorted words that were related to the category “healthy” (e.g., apple, yogurt) with a lever pull (an approach movement) and words related to the category “tasty” (e.g., cookie, fries) with a lever push (an avoidance movement). When participants approached healthy items and avoided tasty items during the sorting task, they subsequently chose more frequently healthy than fatty food when given a choice than a comparison group with the reversed movement assignment (Fishbach & Shah, 2006).

Most impressively, AAT was shown to affect consumptive behaviors involved in alcohol addiction. A seminal study (Wiers, Rinck, Kordts, Houben, & Strack, 2010) trained 42 hazardous drinkers to avoid alcohol-related pictures with a lever push and to approach soft drinks with a lever pull. Results showed less actual beer consumption in a subsequent test-and-rate task among the participants trained to avoid alcohol as compared with controls who were trained to approach alcohol. However, this effect only showed up in a subsample of heavy drinkers for which the AAT procedure proved effective. A subsequent study used a similar training procedure for a treatment of a clinical sample of 214 alcoholic inpatients (Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011). Training the patients to avoid alcohol pictures and to approach picture in four training sessions changed alcohol-approach associations (as indexed with an Implicit Association Test) and reduced self-reported subjective craving relative to control conditions with no or sham training (with no contingency between alcohol pictures and lever responses). Notably, patients in the training group showed 13% less relapse one year after the treatment, suggesting that AAT procedures can have a long-lasting effect on consumptive behaviors. Eberl and colleagues (2013) replicated this finding with a different clinical sample (475 alcohol-dependent patients) and obtained evidence from a moderation analysis that the training effect on the treatment outcome was mediated by a change in the approach bias elicited by the alcohol stimuli (but see also Snelleman, Schoenmakers Tim, & Mheen, 2015). Note, however, that the alcoholic inpatients in these studies received additional therapy; the finding of a long-lasting effect must hence be interpreted with some caution. Sharbanee and colleagues (2014) examined the mechanisms underlying effects of AAT procedures on alcohol consumption using a sample of 74 undergraduate social drinkers. Specifically, they examined whether an effect of AAT procedures on alcohol consumption was mediated by a change in action tendencies (indexed by an approach-avoidance movement task with no contingency between alcohol content and lever movements) or by changes in selective attention to alcoholic beverages (measured with a selective attention task). For instance, participants could have learned to better ignore alcohol-related stimuli during the training phase relative to a control group with no training. Results showed that participants consumed less beer in a test-and-rate task following a training to avoid alcohol relative to an approach training. Importantly, this effect was mediated by a training-induced change in action tendencies, while changes in selective attention had no effect. These results fit with the idea that alcohol AAT procedures diminish beer consumption by a reduction of a habitual approach bias to alcohol (Wiers et al., 2013). However, other studies found no relationship to an alcohol-approach bias (Janssen, Larsen, Vollebergh, & Wiers, 2015; van Hemel-Ruiter, de Jong, & Wiers, 2011) and one study with alcohol-dependent patients found even less relapse when there was a strong alcohol-approach bias (Spruyt et al., 2013). Thus, the relationship between behavioral measures of an approach bias to alcohol-related stimuli and alcohol intake is not clear.

Additional research suggest that AAT procedures can also influence preferences towards non-alcoholic beverages. Participants in one study (Zogmaister, Perugini, & Richetin, 2016) played a video game in which they repeatedly approached one of two juices by dragging it towards themselves with a corresponding movement of the computer mouse (involving an arm flexion), while avoiding the other. Both beverages were novel and thus not known by the participants. Following sufficient training, there was more implicit liking of the approached drink (measured with an IAT), which increased with the measured participants' thirst. There was also a corresponding change in explicit preference ratings but, curiously, this effect was negatively related to participants' thirst. In sum, this research shows that AAT procedures can also change preference towards novel drinks in implicit and explicit preference tests. Note, however, that this research did not include a consumption test of the drinks.

Considering the demonstrated impact of AAT on clinically relevant addiction behavior towards alcoholic soft drinks and the research showing potential effects on soft drink consumption in general, AAT procedures seem to offer a powerful tool to reduce sugary drink consumption and its health consequences which circumvents issues with traditional educational campaigns. Therefore, research is required that provides direct evidence that AAT can affect the consumption of unhealthy drinks.

The present research

The aim of the present research was to demonstrate effects of AAT procedures on consumption of sugary soft drinks. Participants were recruited for consumer research in which they were to test-and-rate two lemonades. The training task was to approach one of two colored lemonades with a lever pull and to avoid the other with a lever push. After sufficient training, they were given an opportunity to taste both lemonades. This consumption test was our main outcome measure. However, we also included explicit and implicit preference measures of the lemonades. In line with the research reviewed above, we hypothesized that participants would consume less of the lemonade that was repeatedly avoided during the training (the avoided drink) relative to the lemonade that was repeatedly approached during the training (the approached drink). We also expected that the avoided drinks would be liked less in an implicit measure (affective priming task) and in an explicit measure (preference rating). Note that a direct within-subject comparison of an approach condition with an avoidance condition should maximize effect sizes in comparison to other control conditions (e.g., a sham training). Furthermore, the potential for problematic confounding variables (such as systematic differences in prior food experiences or drinking habits) is strongly reduced in a design that compares two sugary lemonades as targets for AAT. We therefore planned with sample sizes to have sufficient statistical power for the detection of a medium-sized training effect and stronger. Such an effect size would be consistent with the use of AAT protocols for clinical intervention and applied settings, as it is unlikely that smaller effects would have a measurable impact on outcomes under such circumstances (Wiers et al., 2013).

Experiment 1

Participants were trained to repeatedly approach one lemonade and to avoid the other lemonade based on their color. Task instructions for the training were to respond to a glass filled with red or yellow lemonade by either pulling a joystick lever towards the body (approach) or away from the body (avoidance). A zooming effect (i.e., drawing the lemonade closer or far away on the screen) was not introduced as response effect because a previous study suggested that visually approaching stimuli are appraised less positively (or more negatively) relative to receding or static stimuli (Hsee, Tu, Lu, & Ruan, 2014). However, we randomly mixed in response trials in which participants had to respond to the words ‘towards’ and ‘away’ with a corresponding lever movement. This intermixing was highly effective in previous research to disambiguate the reference point (here: the participant’s body) and the approach- versus avoidance-related meanings of the action (Eder & Rothermund, 2008). Following the training, participants were asked to taste both lemonades. Then, implicit and explicit liking measures were presented in counterbalanced order.

Method

Participants

A total of 65 volunteers (49 female, M_{age} = 27.5, *SD* = 8.9) from the Würzburg area participated in the experiment. We planned to collect data from a minimum of 45 participants to achieve sufficient statistical power for the detection of a medium-sized training effect (*dz* = 0.50) in a one-tailed test; however, in anticipation of data dropouts, data collection continued for the remainder of the scheduled laboratory time. All experiments were approved by a local ethics committee of the Department of Psychology, University of Würzburg (GZEK 2013-14). All participants for all experiments were paid an effective rate of at least 7€ per hour.

Apparatus and stimuli

Up to four participants completed the study at a time, seated at computerized work stations in individual cubicles. Stimulus presentation and measurement of response latencies were controlled by a software timer with video synchronization (E-Prime 2.0 Professional; Psychology Software Tools, Inc.). A Logitech Attack 3 Joystick was attached to the computer.

Training stimuli for the AAT task were pictures of red and yellow lemonades in a typical drinking glass (250 x 400 pixels). A commercially available lemonade (7UP©) was colored red and yellow with a food colorant. For the consumption test, cooled 7UP© was colored and poured into two drinking glasses (33 cl each).

Target words for the affective priming task were selected from the BAWL-R database (Võ et al., 2009) according to their valence and matched for length and frequency (both $|ts| < 1$). Negative targets with bracketed English translations were: MIES (lousy), TRIST (forlorn), BRUTAL (sadistic), WEHRLOS (defenseless), BANKROTT (bankrupt), KRAFTLOS (feeble), MILITANT (militant), TROSTLOS (cheerless), ENTSETZT (horrified), and VERBOTEN (forbidden); positive targets were: TOLL (swell), SUPER (superb), SONNIG (sunny), PFIFFIG (gutsy), TAKTVOLL (tactful), TOLERANT (tolerant), LUKRATIV (lucrative), GRANDIOS (great), SINNLICH (sensual), and REIZVOLL (attractive).

Procedure

After giving informed consent, participants were informed that the study was a consumer taste test for certain beverages. They were offered as many salty pretzels as they liked, after which they rated their feeling of thirst on a 7-point Likert scale from 1 (not at all thirsty) to 7 (very thirsty), followed by their current mood on another 7-point Likert scale from -3 (very bad) to +3 (very good). Thereafter, participants performed the AAT task.

AAT. Participants responded to pictures of colored lemonades by pulling and pushing a joystick lever. In each trial, participants saw an asterisk in the middle of the screen as a fixation point for 200 ms, followed by a blank screen for 100 ms. Thereafter, they saw the target stimulus (drink or action-label word) and responded as instructed. An error message appeared for 2,000 ms if the response was incorrect or slower than 2,000 ms. The next trial began after 500 ms. If the joystick lever was not centered at start, an error message appeared for 2,000 ms, after which the trial restarted.

The target stimuli were pictures of colored drinks in the majority of trials and action-words in a randomly intermixed subset of trials. For the drink pictures, instructions were to pull the joystick lever in response to one colored drink and push it in response to the other drink (counterbalanced assignment). In trials with action-words, participants responded to the German word “HIN” (towards) and “WEG” (away) with a corresponding lever movement towards and away from the body (cf. Eder & Rothermund, 2008). For both trial types, instructions were to respond as quickly and correct as possible. In each block, there were 6 trials with action-labels and 24 trials with drink pictures. Participants received feedback on their average response time and number of errors after each block. The entire procedure was repeated for 6 blocks.

Consumption test. Next, the participant was asked to try as much of both of the lemonades they had seen during the AAT task as they wanted and informed that their opinions on the lemonades would be asked later in the experiment. The experimenter brought one glass of red and one glass of yellow lemonade, each filled with 33 cl liquid. After the participant had finished tasting the lemonades, the experimenter unobtrusively weighed (in gram) how much the participant had drunk of each.

Thereafter, participants completed the affective priming task and answered an explicit preference item asking what drink they liked more. The order of the implicit and explicit tasks was counterbalanced across participants.

Affective priming task. Implicit preferences of drinks were measured using an affective priming procedure (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Participants categorized target words as positive or negative using the “A” and “L” keys on the keyboard (counterbalanced assignment). Instructions were to categorize the words as fast and as accurately as possible. They were also told that before each word, a picture would appear, but that this picture was irrelevant and could be ignored. Each trial began with a fixation cross appearing on the screen for 750 ms, followed by one of the drink pictures from the AAT task for 200 ms. Thereafter, the screen was blank for 50 ms before the target word appeared. An error message appeared for 2,000 ms if the response was incorrect or slower than 2,000 ms. After 500 ms, the next trial was initiated. Participants completed two blocks of 48 trials each, each containing 12 positive-target and 12 negative-target trials per drink.

Explicit rating. Participants indicated their relative preference between the two colored lemonades using a visual analogue scale ranging from -50 to +50. The position of the two lemonades on either side of the visual analogue scale was randomized.

Finally, participants were asked about their suspicions regarding the experiment’s purpose using an open-answer format.

Results

The significance criterion was set to $p < .05$ for all analyses. Greenhouse-Geisser corrected p values are reported with the original degrees of freedom. Standardized effect sizes (Cohen's d, partial eta-square) are reported when appropriate. Timeouts were coded as errors for all analyses. Depending on the analysis, participants were excluded from analyses if their error rates in relevant tasks exceeded the third quartile of the sample by either 1.5 (simple outlier) or three (far outlier) interquartiles (Tukey, 1977). No participant produced far outlier values in the error rates of the AAT (criterion: 15.3%). For the label trials, the more conservative simple outlier criterion was applied (criterion: 23.6%), as participants with high error rates in these trials arguably did not follow the approach-avoidance instructions for the AAT task (cf. Eder & Rothermund, 2008). This led to the exclusion of seven participants.

Drink color did not interact with the AAT factor in any analyses. Therefore, this factor was not included in the analyses reported below. Drinks were compared according to whether they were paired with an approach- or an avoidance-related response during the training phase (henceforth: approached versus avoided drinks). Raw data underlying the findings can be retrieved at Harvard Dataverse.

Bayesian analysis parameters. To enable statements about evidence in favor of the null hypothesis, Bayesian analyses were additionally conducted using the JASP software package (JASP Team, 2018). The critical value of Bayesian tests is the Bayes factor, which indicates the relative likelihood of the observed data under the alternative hypothesis compared to the null hypothesis (BF_{10}). The larger the value of BF_{10}, the stronger the evidence for the alternative hypothesis: for $BF_{10} = 10$, the observed data are 10 times more likely under the alternative hypothesis than under the null hypothesis. The Bayes factor may also be written as BF_{01}, which is the inverse of BF_{10}. Therefore, a BF_{01} of 10 indicates that the observed data are 10 times more likely under the null hypothesis than under the alternative hypothesis. For the following analyses, Bayes factors will always be reported in “larger is stronger” form, meaning that when the balance of evidence favors the alternative hypothesis, BF_{10} will be reported, but when it favors the null hypothesis, BF_{01} will be reported. Jeffreys (1961) suggests that Bayes factors of higher than 3 in favor of a given hypothesis may be seen as substantial evidence for that hypothesis, whereas higher than 10 may be considered strong, higher than 30 very strong and higher than 100 extreme. Analogously to the null hypothesis significance tests (NHST), one-sided Bayesian t-tests were conducted to assess the effects of the AAT factor. Unless otherwise noted, the default Cauchy prior width of .707 was used in all analyses. Robustness checks for the influence of the prior were conducted using additional prior widths of 1.000 and 1.414. If these priors weakened the evidence compared to the default prior, it is reported with the relevant analysis.

AAT performance. Training performance was very high (proportion of correct responses M = 97.2%, SD = 2.2, range: 90-100%). Correct performance in the label trials was M = 92.7% (SD = 5.2; range: 81-100%).

Drink consumption. The amount of drink consumed (measured in grams) was compared on the AAT factor using paired-samples t-tests. There were no effects on the consumption of approached drinks (M = 96 g, SD = 56) and avoided drinks (M = 96 g, SD = 52), $t(57) = 0.05$, $p = .482$. In line with this result, the paired-sample Bayesian t-tests for consumption showed substantial evidence for the null hypothesis ($BF_{01} = 6.71$).

Implicit preference test. The first four trials were discarded as practice. Three participants with high error rates were excluded based on a far outlier criterion (criterion: 20.7%). In the remaining data, trials with no or incorrect responses (5%), anticipations with a RT below 100 ms and/or above their personal simple outlier criterion (5.9%) were eliminated before RT analysis. Outlier removal did not alter the pattern of results in this and in subsequent experiments. Average reaction times to positive targets following a specific drink were subtracted from average reaction times to negative targets for the same drink to form an index of implicit positivity for that drink (with positive values indicating more positivity). An analogous score was computed for the error rates by subtracting errors to positive targets from errors to negative targets. Order of implicit and explicit liking measures had no effect on the results (largest $F = 1.13$, $p > .20$) and was therefore not included in the analyses reported below.

Implicit positivity scores were not different for approached versus avoided drinks in the RT measure, $t(54) = 0.20$, $p = .47$, and in the error measure, $t(54) = -0.41$, $p = .68$ (see Table 1 for the descriptive statistics). Paired-sample Bayesian t-tests for implicit positivity showed substantial evidence for the null hypothesis for the RT measure ($BF_{01} = 5.78$) and the error measure ($BF_{01} = 9.08$).

Table 1

Means of reaction times (in ms) and error rates (in percent) in the affective priming task of Experiment 1. Standard deviation is shown in parentheses.

Target	Approach-trained prime		Avoidance-trained prime	
Positive	594 ms (81)	5.7% (6.2)	594 ms (85)	6.0% (5.5)
Negative	613 ms (85)	3.9% (4.3)	612 ms (87)	4.6% (4.8)

Explicit preference test. VAS ratings were scored that positive values indicate a preference for the approached drink. Order of the explicit and liking measure had no effect on the VAS ratings. A preference for the approach-trained drink was analyzed using a one-sample t-test against a value of 0, revealing no training effect ($M = 0.6$, $SD = 26.0$), $t(57) = 0.18$, $p = .860$. In line with the frequentist analysis, a one-sample Bayesian t-test produced substantial evidence for the null hypothesis ($BF_{01} = 6.04$).

Regression analyses. Participants rated their thirst with a medium value ($M = 4.3$, $SD = 1.4$) on a scale ranging from 1 (not at all thirsty) to 7 (very thirsty). For an assessment of a possible moderation by thirst (see Zogmaister et al., 2016), separate regression analyses were performed with mean-centered thirst as predictor variable and difference scores for drink consumption (consumption of the approached drink minus avoided drink collapsed across both drink sessions), implicit positivity scores (RT score for the approached drink minus RT score for the avoided drink), and relative VAS ratings, respectively, as criterion variables. These analyses produced no significant result (drink consumption: β = -.10; implicit positivity: β = .13; explicit VAS rating: β = .001; all |*ts*| < 1).

Discussion

The results are clear-cut. The AAT procedure had no effects on the outcome measures, and Bayesian analyses provided substantial evidence for the null hypothesis. One possible explanation for the null effects is that approach and avoidance goals were not sufficiently induced during the training by an intermixing of a few action-word trials. In addition, the emphasis on a product test might have reduced participants' motivation to consume the lemonade of their preference, because of a belief that they must consume both drinks in equal amounts to come up with an unbiased judgment. Therefore, we changed these procedures for a second experiment.

Experiment 2

In Experiment, the drink picture on the computer screen consistently zoomed towards the participant following a lever pull and away from the participant after a lever pull. With this zooming effect, we hoped for a clear disambiguation of the action meanings in terms of approach and avoidance (see Rinck & Becker, 2007). Furthermore, task instructions for the consumption test were adapted to suggest less of a consumer focus and more of a focus on personal preferences. The consumption test was repeated after a second training session for a more reliable outcome measure. In addition, a behavioral assessment task was included at the end that probed for a systematic changes in automatic behavioral tendencies in line with the AAT procedure. This behavior test was included to evaluate the effectiveness of the training procedure to produce a longer-lasting change in approach-avoidance tendencies.

Method

Participants

Participants were 36 volunteers from the Würzburg area (27 female, M_{age} = 28.8, SD_{age} = 10.3). Sample size was planned to detect a medium-sized effect (*n* = 27), but data collection was again continued to maximize scheduled laboratory use.

Apparatus, stimuli, and procedure

Apparatus and stimuli were the same as in Experiment 1. The procedure for the AAT task was the same except for the following changes: Instead of using trials with action-labels, a zooming effect was added to the responses to establish a reference to approach and avoidance (Rinck & Becker, 2007). The zooming lasted 300ms. Furthermore, two pictures showing half-full glasses were added to the picture set of drinks. Participants completed 5 blocks of 30 trials each before consumption of the lemonades for the first time, and an additional 5 blocks after the first tasting. A second consumption test was provided after this session. In both consumption tests, the drinking glass was refilled when the glass was emptied. The lemonades in this experiment were 7UP© and Sprite©, which have a similar but distinguishable lemon-lime flavor. The lemonades were colored red and yellow as in Experiment 1. The slightly different taste of a second lemonade should enhance the plausibility of a preference test.

The behavioral assessment task was identical with the AAT task, except that the relevant response cue was now whether the glass was full or half-full. Accordingly, the SR mapping for the behavior test was orthogonal to the SR mapping that was established for the AAT task, leading to SR combinations that were congruent or incongruent with the learned response for the AAT task. Half of the trials were congruent, half incongruent; of these, half required a pull response, half a push response. The behavioral assessment had two blocks with 24 trials each. Lever movements were without a zooming effect in this task. Participants completed the assessment task as a final measure immediately after the liking measures.

Results

The same data-analytic procedures were employed as in Experiment 1.

AAT performance. Correct performance was high (*M* = 98.1%, *SD* = 1.5, range: 95-100%).

Drink consumption. The amount of drink consumed was analyzed with an analysis of variance (ANOVA) with AAT (approached vs avoided drink) and drink session (first vs second) as factors. The main effect of drink session was significant, $F(1, 35) = 6.17$, $p < .05$, $\eta^2 = .15$. Participants drank more in the first session. More important, the main effect of AAT (approached drink *M* = 65g, avoided drink *M* = 70g), and the interaction between both factors were not significant (both *F*s < 1).

A Bayesian repeated measures ANOVA corroborated the results from the frequentist analyses. The model that received the most support against the Null model is the model stating a main effect of drink session (*BF10* = 13.33). In contrast, evidence for the null model was strongest for the model claiming a main effect of AAT (*BF01* = 4.01). In short, odds favor a model that the training procedure did not affect the consumption of the drinks.

Implicit preference test. Performance in the affective priming task was analyzed in the same way as in Experiment 1. From the RT data, 7.1% were removed due to errors and 5.5% as simple RT outliers according to Tukey (1977). Implicit positivity RT scores were subjected to a mixed ANOVA with AAT as within-factor and the order of the liking measures (implicit first versus explicit first) as between-factor. The main effects of measure order and AAT were not significant (with both *F*s < 1). The two-way interaction achieved significance, $F(1, 34) = 8.48$, $p = .006$, $\eta^2 = .20$. As shown in Table 2, implicit positivity scores were higher for the approached drink relative to the avoided drink when the explicit measure was presented first, and vice versa when the implicit measure came first. One-sided follow-up comparisons showed that the effect in the explicit-first order condition was significant, $t(18) = 2.39$, $p = .01$, while the (reversed) effect in the implicit-first order condition was not, $t(16) = -1.78$, $p = .18$. An analogous ANOVA of the error rates produced no significant result (all *F*s < 1).

In the corresponding Bayesian repeated measures ANOVA of the RTs, the evidence favors the null model over a model including an AAT effect term for the implicit positivity scores ($BF01 = 3.98$). In addition, the Bayesian analysis revealed only anecdotal evidence for a model including an interaction effect between AAT and measure order in comparison to a null model including only the main effects of AAT and measure order ($BF10 = 2.39$). Thus, evidence for an AAT effect in the explicit-first condition is spurious at best. Bayesian analyses of the error rates produced substantial evidence favoring the null model over all other factorial models (smallest $BF01 = 3.07$).

Table 2

Means of reaction times (in ms) and error rates (in percent) as a function of measure order in the affective priming task of Experiment 2. Standard deviation is shown in parentheses.

Measure order	Target	Approach-trained prime		Avoidance-trained prime	
Implicit first	Positive	613 ms (90)	7.8% (6.7)	605 ms (83)	6.5% (5.1)
	Negative	611ms (68)	8.5% (7.7)	625 ms (101)	5.7% (7.1)
Explicit first	Positive	580 ms (66)	6.2% (5.7)	592 ms (74)	8.2% (8.4)
	Negative	608 ms (67)	6.7% (6.2)	596 ms (76)	7.4% (6.2)

Explicit preference test. VAS ratings were scored in the same way as for Experiment 1. The order of the explicit and implicit liking measures had no effect on the VAS rating ($F < 1$). A one-sample t-test against zero showed no significant difference in the relative rating of the approached versus the avoided drink ($M = -7.0$, $SD = 29.3$), $t(35) = -1.44$, $p = .16$. In addition, a one-sample Bayesian t-test produced strong evidence for the null hypothesis ($BF_{01} = 12.59$).

Behavior assessment task. In analyses of the behavior assessment task, the data from one participant who responded at chance (46% errors) was removed. The first two trials were dropped as task practice. Trials with errors (3.0%) and RT simple outliers (6.3%) were additionally removed before RT analyses. Performance (RT, errors) was analyzed depending on whether the color association established with the AAT procedure afforded the same response (congruent trials) or a different response than the one instructed for the categorization task (half-full vs. full). One-sided paired-samples t-tests showed no effect of congruence with trained responses in RTs (congruent M = 498 ms, incongruent M = 495 ms); $t < 1$, and error frequencies (congruent M = 2.4%, incongruent M = 3.5%), $t(34) = -1.46$, $p = .07$.

Regression analyses. Participants indicated moderate thirst (M = 3.7, SD = 1.2). Separate regression analyses with mean-centered thirst as predictor variable and difference scores for drink consumption (consumption of the approached drink minus avoided drink collapsed across both drink sessions, $\beta = -.03$), implicit positivity scores (RT score for the approached drink minus RT score for the avoided drink, $\beta = .01$), and relative VAS ratings ($\beta = -.20$), respectively, as criterion variables produced no significant results (largest $|t| = 1.18$). Furthermore, analogous regression analyses were carried out with scores of the behavior assessment task (RT in incongruent trials minus RT in congruent trials) as predictor variable. These regression analyses analogously produced no significant result (drink consumption: $\beta = .07$; implicit positivity: $\beta = -.27$; explicit VAS rating: $\beta = .12$; largest $|t| = 1.62$).

Discussion

We again found no training effect on the outcome measure despite a change in the cover story and despite the introduction of a zooming effect. There was a small training effect in the implicit measure when the explicit rating was presented first, but Bayesian analyses provided only anecdotal evidence for this interaction effect. Even more important, the behavioral assessment task provided no evidence for a training-induced change in action tendencies. This result suggests that the AAT procedure was not effective in inducing an approach (or avoidance) bias in line with the training schedule. Note, however, that the behavioral assessment was presented as a final measure after the outcome measures. Therefore, it is possible that the AAT training was effective but that the training-induced change was subjected to temporal decay and/or interference by the measurement tasks. Experiment 3 was therefore designed to evaluate this possibility.

Experiment 3

In Experiment 3, participants alternated between AAT training blocks and blocks of the behavior assessment task. Previous research showed that performing intentional approach and avoidance actions in response to positive and negative stimuli affects the unintended activation of approach and avoidance tendencies in an intermixed task for which stimulus valence was irrelevant (Eder, Rothermund, & Proctor, 2010). By intermixing blocks of AAT trials with blocks of the behavior assessment task, we were able to analogously track changes in automatic action tendencies induced by the AAT schedule.

Method

Participants

For sample size planning, we used a sequential Bayesian hypothesis testing rule (see Schönbrodt, Wagenmakers, Zehetleitner, & Perugini, 2015). The goal was to continue collecting data until the Bayes factors on all tests of interest exceeded 10 in favor of either the null or alternative hypothesis. Importantly, this sampling procedure is not dependent on correct effect size guesses in a priori power analyses and it permits unlimited multiple testing until the predefined level of evidence was reached. The first Bayesian analyses were conducted at 41 participants. Thereafter, sequential analyses were calculated after every week of data collection. Due to economic reasons, the criterion could not be achieved and data collection was terminated at 114 participants (85 female, M_{age} = 26.1, SD_{age} = 8.4). Frequentist analyses were exclusively performed with this final sample.

Apparatus, stimuli, and procedure

Apparatus, stimuli, and procedures were the same as in Experiment 2. However, this time participants completed 12 AAT training blocks of 20 trials each in which they responded to the color of the drinks. These blocks alternated predictably with blocks of the behavior assessment task in which responses were made to the fill status of the drinks (half-full versus full). This task was the same as in Experiment 2, except that a zooming effect was added to correct responses. The assessment task consisted of 12 blocks with 20 trials each (10 congruent, 10 incongruent trials). In contrast to Experiment 2, only one consumption test was implemented following the training phase. Furthermore, the explicit preference judgment was now a 5-point Likert scale with values on either side indicating a strong preference for the selected drink and the middle point indicating no preference. The allocation of the yellow and red drinks to the left and right sides of the scale was random. All other aspects of the procedure were identical with Experiment 2.

Results

The same data analytic procedures were employed as in Experiment 2. Two participants were removed from analyses as far outliers in AAT performance (criterion: 9.2% errors).

AAT performance. Correct performance in the AAT was very high (*M* = 97.9%, *SD* = 1.8%, range: 92-100%).

Behavior assessment test. Trials with errors (2.4%) or RT simple outliers (5.3%) were removed before RT analyses. Performance was analyzed depending on whether the instructed responses was congruent or incongruent with the trained response in the AAT blocks. One-sided paired-samples t-tests provided clear evidence for a congruence effect in the RT measure (congruent *M* = 501 ms, incongruent *M* = 505 ms), $t(111) = 3.38$, $p < .001$, $d_z = 0.32$, and in the error measure (congruent *M* = 1.9%, incongruent *M* = 2.9%), $t(111) = 4.92$, $p < .001$, $d_z = 0.46$.

Drink consumption. The amount of drink consumed was compared with one-sided t-tests. Participants consumed approximately the same amount of the approached drink (*M* = 88g) and the avoided drink (*M* = 92g), $t < 1$. In line with this result, a Bayesian paired-samples t-test provided strong evidence for the null model (*BF01* = 16.88).

Implicit preference test. One data set had to be removed because of an excessive error rate (23%). In addition, 5.1% of the reaction times were removed due to errors and 5.3% were simple RT outliers. Order of implicit and explicit liking measures had no effect on the results (largest $F[1,110] = 1.20$). Paired-samples t-tests revealed that implicit positivity scores were not different for approached versus avoided drinks in the RT measure ($t < 1$) and in the error measure, $t(110) = -1.63$, $p = .11$ (for means and SDs see Table 3). Corresponding Bayesian t-tests showed substantial evidence for the null model in the RT measure ($BF_{01} = 4.27$) and anecdotal evidence for a null effect in the error measure ($BF_{01} = 1.47$).

Table 3

Means of reaction times (in ms) and error rates (in percent) in the affective priming task of Experiment 3. Standard deviation is shown in parentheses.

Target	Approach-trained prime		Avoidance-trained prime	
Positive	637 ms (79)	4.1% (5.3)	642 ms (82)	5.1% (5.3)
Negative	660 ms (90)	6.2% (6.1)	662 ms (94)	5.7% (6.5)

Explicit preference test. Ratings were recoded that positive values indicate a preference for the approached drink. Order of the explicit and implicit measure had no effect on the ratings, $F(1,111) = 1.85$, $p > .15$. A one-sample t-test against zero showed no effect of AAT ($M = 0.2$, $SD = 1.1$, $t < 1$), and the corresponding Bayesian t-test revealed substantial evidence for the null model ($BF_{01} = 8.27$).

Regression analyses. Participants indicated moderate thirst ($M = 4.4$, $SD = 1.3$). As for Experiment 2, outcome measures were regressed onto thirst ratings and onto congruency effects in the behavior assessment task (indexing training-induced changes in automatic action tendencies). Regression analyses with mean-centered thirst as predictor variable showed no significant results for drink consumption (β = .10), implicit positivity scores (β = .03), and explicit preference judgments (β = .12), largest $|t|$ = 1.24, all *p*s > .20. Regression analyses with congruency effects in the behavior assessment test as predictor variable analogously produced no significant result (drink consumption: β = -.01; implicit positivity: β = -.08; explicit preference: β = .13; largest $|t|$ = 1.35, all *p*s > .18).

Discussion

Results were again in line with a null model claiming no training effect. Notably, this result was obtained although the behavior test provided clear evidence that action tendencies were changed by the training. Due to the large sample size (*n* = 114), statistical power this time was even good enough for a detection of small effects (*dz* = 0.24 with 1- β = .80). Furthermore, Bayesian analyses consistently supported a null model on either outcome measure. Thus, it appears that our AAT procedures, even though effective, did not affect drink consumption and/or preference scores.

Experiment 4

The design of the experiments described so far was optimized to detect a training effect on drink consumption. Immediately after training, participants had an opportunity to consume the drinks, followed by implicit and explicit liking measures in counterbalanced order. With this measurement order, the consumption of the drinks might have systematically affected the subsequent liking measure. For instance, participants could have adjusted their drink attitudes after they have noticed that both lemonades have a very similar taste. Accordingly, presenting the consumption test first may have overridden possible training effects on the liking measures (see Zogmaister et al., 2016).

In Experiment 4, we tested this possibility with a variation of the measurement order: attitude measures were presented either before or after the consumption test. In addition, participants were asked to provide more specific explicit ratings of the drinks. Specifically, they were asked to rate the taste, healthiness, and the drinks’ ability to quench thirst. We reasoned that the drink consumption would inform subsequent judgments about the drinks, especially ratings of their taste and thirst-quenching ability. Importantly, this information was not available in the condition when the attitude measures were presented first. Accordingly, we hypothesized that the AAT procedure should bias participants' implicit and/or explicit attitudes about the drinks before the consumption test but not after the consumption of the drinks.

Method

Participants

The study followed a 2x2 mixed design with the between-factor measurement order (consumption first vs. attitudes first) and the within factor AAT (approach-trained drink vs. avoidance-trained drink). We planned with a total sample size of 90 (sensitive to an interaction effect of $f \geq .149$ with at least 80% statistical power). The final sample consisted of 97 participants (67 female, $M_{age} = 26.9$, $SD_{age} = 8.4$; age data of one participant was lost due to a technical error). The data collection stopping rule was similar to Experiments 1 & 2.

Apparatus, stimuli, and procedure

The AAT was the same as in Experiment 3. Participants completed 12 blocks of 20 trials each. As no behavior assessment test was implemented in this experiment, participants completed the training without interruptions. Following the training, half of the sample consumed the drinks first and completed the attitude measures second (the consumption first group). The other half completed the measures in a reversed order (the attitudes first group). For both groups, the explicit attitudes towards the drinks were measured separately on three 7-point Likert scale items: taste (from -3 "very unpleasant" to 3 "very pleasant"), healthiness (from -3 "very unhealthy" to 3 "very healthy") and ability to quench thirst (from -3 "hardly able to quench thirst" to 3 "very able to quench thirst"). The order of the implicit and explicit attitude measures was counterbalanced across participants. All other aspects of the procedure were identical with Experiment 1.

Results

Pre-analyses were the same as for Experiment 1. One participant in the consumption first group did not drink any soda due to Ramadan; her data were dropped. One additional participant had to be excluded due to an unusually high error rate in the AAT task (criterion: 9.2%).

AAT performance. As in the experiments before, participants had no difficulties with the training task (correct performance M = 97.9%, SD = 1.7, range: 91-100%).

Drink consumption. The amount of drink consumed was analyzed with a mixed ANOVA with approached versus avoided drink (AAT) as within-factor and measurement order (consumption first vs attitudes first) as between-factor. Results showed that participants consumed approximately the same amount of the approached drink (M = 79g) and the avoided drink (M = 85g), $F(1,93) = 1.47$, $p > .20$, irrespective of measurement order ($F < 1$). The main effect of measurement order was significant, $F(1,93) = 5.69$, $p < .05$, $\eta^2 = .058$. Participants drank more when the consumption test was presented first (M = 96g) compared to when the attitude measures came first (M = 69g).

A corresponding mixed Bayesian ANOVA provided anecdotal evidence for the model including only a main effect of measurement order (*BF10* = 2.17). More importantly, the analysis showed evidence against inclusion of the training main effect (*BF01* = 4.08) and the interaction effect (*BF01* = 6.94).

Implicit attitude measure. For analyses of the affective priming task, one data set had to be removed because of an excessive error rate (far outlier criterion: 26.1%). In addition, 5.7% of the affective priming trials were eliminated as errors; of the correct trials, 0.6% were eliminated as anticipations and 4.9% as simple RT outliers. Implicit positivity RT scores were subjected to a mixed ANOVA with AAT (approached vs avoided drink) as within-factor and measurement order (consumption first vs attitudes first) and order of the attitude measures (implicit measure first vs. explicit measure first) as between-factors. No main effects or interactions approached significance (largest $F[1,90] = 1.51$), Descriptive statistics can be seen in Table 4.

An analogous ANOVA of the error rates produced a significant interaction between AAT and measurement order, $F(1, 90) = 4.25$, $p < .05$, $\eta^2 = .045$. Implicit positivity scores were larger for the approached drink (*M* = 1.4%) relative to the avoided drink (*M* = -2.0%) when the consumption test came first, while scores for the approached drink were reduced (*M* = 0.1%) relative to the avoided drink (*M* = 1.2%) when the attitude measure was presented first. Follow-up comparisons showed that enhanced implicit liking of the approached drink after consumption of the drinks was significant, $t(44) = 2.31$, $p < .05$, $dz = 0.34$, while a lowered liking of the approached drink before drink consumption was not ($|t| < 1$).

A Bayesian ANOVA of the implicit positivity scores (RT measure) with AAT as within-factor and measurement order and order of the attitude measures as between-factor did not support inclusion of a main effect of training (*BF01* = 13.70) or any interaction term including training (smallest *BF01* = 50.00). A Bayesian analysis of the error rates produced similar results (main effect: *BF01* = 6.67, smallest value for interactions: *BF01* = 4.52).

Table 4

Means of reaction times (in ms) and error rates (in percent) as a function of measure order in the affective priming task of Experiment 4. Standard deviation is shown in parentheses.

Measurement order	Target	Approach-trained prime		Avoidance-trained prime	
Consumption first	Positive	625 ms (92)	4.7% (6.2)	621 ms (85)	8.1% (7.6)
	Negative	645 ms (92)	6.1% (6.0)	639 ms (93)	6.0% (6.0)
Attitudes first	Positive	628 ms (100)	5.2% (5.9)	623 ms (96)	4.5% (4.9)
	Negative	644 ms (93)	5.3% (6.1)	645 ms (101)	5.8% (6.4)

Explicit rating measures. Ratings of taste, healthiness, and thirst-quenching for the approach-trained and the avoidance-trained drinks were highly interrelated with Cronbach's alpha = .66 and .70, respectively. Furthermore, order of the explicit and liking measures had no effect on the results; this factor was therefore not included in the analyses reported next. Table 5 displays means with standard deviations for each rating item.

A mixed ANOVA of taste ratings with AAT as within-factor and measurement order as between-subjects factor produced no significant result (largest F[1,93] = 1.43). Ratings of healthiness indicated that the lemonades were generally appraised as unhealthy. However, the ANOVA revealed no significant effect of AAT and/or measurement order on the ratings scores (all Fs < 1). An analogous ANOVA of the thirst-quenching rating scores produced a main effect of measurement order, $F(1, 93) = 7.22$, $p < .01$, $\eta^2 = .072$. Drinks were rated as less thirst-quenching following their consumption. More importantly, AAT had no effect on the rating scores, $F(1, 93) = 1.52$, $p = .22$, irrespective of whether the drinks were consumed before or after the rating ($F < 1$).

Table 5

Mean rating (with standard deviation) of the taste, healthiness, and ability to quench thirst as a function of measurement order in Experiment 4.

Measurement order	Drink	Taste	Healthiness	Thirst-quenching
Consumption first	approached	0.6 (1.6)	-1.7 (1.1)	-0.2 (1.5)
	avoided	0.5 (1.6)	-1.6 (1.3)	-0.1 (1.6)
Consumption last	approached	0.8 (1.6)	-1.6 (1.2)	0.6 (1.4)
	avoided	0.4 (1.5)	-1.5 (1.3)	0.7 (1.4)

Bayesian analyses of the ratings scores corroborated the NHST analyses. Mixed Bayesian ANOVAs with AAT and measurement order as factors did not support inclusion of AAT main effects or the two-way interaction on ratings of taste (main effect: *BF01* = 4.59, interaction effect: *BF01* = 21.74), healthiness (main effect: *BF01* = 7.94, interaction effect: *BF01* = 31.25), and ability to quench thirst (main effect: *BF01* = 3.98, interaction effect: *BF01* = 5.95).

Regression analyses. Outcome measures (with positive scores indicating a preference for the approached drink) were regressed on mean-centered thirst ratings ($M = 4.3$, $SD = 1.4$). These analyses produced a significant result for the consumption measure ($\beta = -0.28$), $t(93) = -2.85$, $p < .01$. Opposite to expectations, thirsty participants tended to consume more of the avoided drink. Standardized regression coefficients for difference scores of implicit positivity ($\beta = -0.03$), taste ratings ($\beta = -0.16$), healthiness ratings ($\beta = 0.02$), and thirst-quenching ratings ($\beta = -0.14$) were not significant (largest $|t| = 1.53$, all *p*s > .10)

Discussion

Experiment 4 examined whether the consumption of the drinks immediately after the training systematically affected participants' attitudes of the drinks as measured with an implicit task and an explicit rating task. The results clearly showed no moderation by this procedure. NHST analyses found weak evidence for a training-induced change in implicit preference scores (measured with an affective priming task) when the drinks were consumed before the implicit measure. However, the corresponding interaction effect was significant only in the error measure, and Bayesian analyses supported a null model of this interaction effect more. Therefore, this result should be interpreted with caution. For the explicit measures, we obtained no evidence for a training effect, or for a moderation by drink consumption. Thus, it appears that consumption of the drinks was not the reason why we have found no training effects in attitude measure.

Mega-analysis

It is possible that our AAT procedures had very subtle effects on the outcome measures that could not be detected in the individual experiments. Therefore, we performed a mega-analysis that jointly analyzed the pooled data from all of the experiments. With the increased sample size (n = 300), NHST analyses have sufficient statistical power (1- β > .95) to detect even small training effects (dz >= 0.2). In addition, Bayesian analyses benefit from a large sample size because the posterior distribution is less influenced by the analyst's selection of priors. As a consequence, results of Bayesian analyses become more independent of the particular choice of the a priori beliefs and could be interpreted with more confidence. For a mega-analysis, difference scores were computed for each outcome measure (consumption, implicit preference [RT], explicit preference) in a way that positive values indicate more consumption and more preference for the approached drink relative to the avoided drink. The explicit ratings in Experiment 4 were pooled for a single index. Difference scores were then standardized by dividing them through their standard deviations (resulting in a d-measure), and the standardized values were then analyzed with frequentist and Bayesian tests.

NHST analyses with one-tailed t-tests against zero produced no significant result for drink consumption (M = -4.5, SD = 48.0), t(299) = -1.63, implicit preference test (M = 1.4 ms, SD = 47.3), t(295) = 0.48, and explicit preferences (M = -0.02, SD = 1.00), t(299) = -0.28. Note that the means for the consumption test and explicit rating item had negative signs indicating a more favorable disposition towards the avoided drinks. In line with these results, corresponding one-tailed Bayesian t-tests against zero produced strong evidence for a null model of drink consumption (*BF01* = 39.4), implicit preferences (*BF01* = 10.0), and explicit preferences (*BF01* = 19.1).

General discussion

Consumption of sugary soft drinks has exploded in Western societies over the last four decades, which is problematic given its relation to several medical disorders (Mensink u. a., 2018). The present study examined whether consumption of soft drinks could be changed with a relatively new behavioral training method referred to as 'approach-avoidance training' (AAT). Four experiments examined whether explicit training of approach and avoidance responses to soda drinks affects the consumption of soft drinks immediately after the training and implicit and explicit preferences towards the drinks. In standard NHST analyses, no training effects were found on either measure. This result was obtained although statistical power of standard NHST analyses was sufficient to detect a small-to-medium sized effect. In addition, Bayesian analyses consistently provided more support for a null model than for the alternative model. Thus, the conclusion is warranted that our AAT procedure was not effective in producing a change in soft drink consumption and in implicit and explicit attitudes towards the drinks.

Before we discuss possible reasons for this result, we want to emphasize that the present research deviated in several ways from previous AAT studies. Most published AAT studies examined training effects on alcohol consumption with alcohol-dependent or at-risk drinkers. Thus, there was a systematic difference in the populations under study. Furthermore, our experimental procedures differed from previous AAT studies in several ways. In the following sections, we will discuss the most striking procedural differences in more detail.

Relevant/irrelevant AAT task. In the present research, the AAT instructions explicitly assigned approach- and avoidance-related movement to the lemonades, which means that the lemonades were relevant stimuli for the training. This procedure differs from most previous AAT studies in which responses were made to an irrelevant feature of alcohol-related stimuli (e.g., picture format). We intentionally decided against an irrelevant AAT task for several reasons. First, based on learning theories, one should expect stronger (associative) learning when attention is directed towards the to-be-trained SR contingencies (Mackintosh, 1975; Pearce & Hall, 1980; for a recent review see Mitchell & Pelley, 2010). In fact, we could not come up with a single learning theory that would predict better learning with a training program that directs attention away from the relevant knowledge. There is a reduced risk for demand characteristics with an implicit training schedule but this should decrease, rather than increase, the likelihood of a training effect. Most important, the theoretical argument for improved training effects with a relevant AAT task is also supported by empirical research. Van Dessel and colleagues found no AAT effects on (explicit and implicit) stimulus evaluations in a subliminal training task but robust effects in a supraliminal AAT task or when participants were explicitly informed about the stimulus-action contingencies (Van Dessel, De Houwer, Roets, & Gast, 2016). Another study of this research group found no AAT effects on implicit and explicit stimulus evaluations when participants had no contingency knowledge of the stimulus-action relationship (Van Dessel, De Houwer, & Gast, 2016). According to this research, a direct instruction of stimulus-action relationships during AAT training should produce stronger training effect or at least effects that are comparable with an extended irrelevant AAT task (see Field, Caren, Fernie, & De Houwer, 2011). As a matter of fact, when Wiers and colleagues (2011) compared extended training conditions with explicit and implicit stimulus-action contingencies in a study, they found no significant differences between both training procedures. Therefore, the use of a relevant AAT task for training is not a plausible explanation of the null results in the present research.

Effectiveness of the AAT procedure. In the present experiments, the number of training trials ranged between 144 trials in Experiment 1, 240 trials in Experiments 3 & 4, and 300 trials in Experiment 2. Previous AAT studies in the alcohol domain often had significantly more training trials but there exists little systematic research on how many training trials are required for a change in AA tendencies. Eberl and colleagues (2014) had 12 training sessions with 160 trials each and found that 6 training sessions were optimal for the prediction of 1-year alcohol relapse. However, other studies obtained training effects with significantly fewer training trials that were in the range of the present study (e.g., Van Dessel et al., 2018). Furthermore, it is plausible that the effectiveness of the AAT procedure is influenced by additional features of the training task (e.g., relevant versus irrelevant training task, contingency awareness, sensitivity of outcome measures, etc.), so that a few training trials could be sufficient. In fact, one study observed an AAT effect even in the complete absence of training when actions were merely instructed and not performed by the participant (Van Dessel, De Houwer, Gast, & Tucker Smith, 2014).

The present research also had behavior assessment tests checking the effectiveness of the training procedure in two of the four experiments. In Experiment 2, a behavior test was presented following the outcome measures at the end of the session. Results did not provide evidence for a training-induced effect on action tendencies, suggesting that the AAT method was not effective in producing a longer-lasting change in AA tendencies. Experiment 3 intermixed blocks of the behavior assessment test with blocks of the AAT task, and obtained clear evidence for a training-induced change in automatic action tendencies. It should be noted that zooming effects disambiguating the lever movements as AA responses were presented only in the behavior test of Experiment 3 but not in Experiment 2, which might explain the reduced sensitivity of the latter test. Another possibility is that the AAT procedure had a strong effect on action tendencies only in the short-term. In fact, one study using a similar task procedure observed nearly-instant changes in automatic AA tendencies when an evaluation task with intentional AA responses to affective stimuli was intermixed with an affective Simon task measuring unintentional AA tendencies towards those stimuli (Eder et al., 2010; see also Eder, 2011). It should be noted that a transient change in AA tendencies, as observed in Experiment 3, should have at least affected the outcomes measures immediately following the training. This was clearly not observed. Furthermore, if the effects of the AAT were short-lived, they would have had to be extremely so to evade detection in our mega-analysis, which should have been sufficiently powered to find even the remnants of a fading effect. Therefore, it is unwarranted to explain away the results with fleeting effects of the training procedure. Instead, our research raises the question of what determines whether AA tendencies affect subsequent evaluations and behavior, as well as what variables might affect the stability of AAT-induced AA tendencies.

Retraining of appetitive reactions. As reviewed in the introduction, most previous studies that were successful in a retraining of consumptive behaviors were carried out with alcohol-dependent drinkers or at-risk drinkers. This population differs from groups of ordinary people. Alcohol dependence is typically characterized by strong appetitive reactions and attentional biases towards alcohol consumption (Field et al., 2008; Townshend & Duka, 2001). Thus, the heavy drinkers in AAT studies presumably had strong preexisting motivations that were modified by the AAT procedures (e.g., Wiers et al., 2010). By contrast, participants in the present experiments exhibited only weak motivations to consume the drinks. Although most people like lemonades and measures were taken to increase participants' thirst before the training (e.g., by eating salty pretzels), this could not induce appetitive reactions of comparable strength to addictive motivations. Furthermore, research showed that heavy drinkers are often ambivalent in their craving for alcohol, with conflicting tendencies to drink alcohol and to avoid excessive drinking (Stritzke, McEvoy, Wheat, Dyer, & French, 2007). Thus, alcohol-related stimuli before training may trigger both a motivational tendency to approach alcohol and a motivational tendency to avoid alcohol abuse. Wiers and colleagues (2011) argued that AAT procedures may exert an effect by increasing the relative accessibility of alcohol-avoid and alcohol-approach associations. With extended training to avoid alcohol-related stimuli, the corresponding alcohol-avoid association is strengthened and a preexisting approach bias is changed to an avoidance bias. A modification of conflicting action tendencies towards alcohol thus could explain why many alcohol AAT studies obtained a training effect (but see also Lindgren et al., 2015; Wiers et al., 2015, for finding no effect). It should be noted, however, that AAT studies in other domains obtained training effects even when they used novel stimuli (e.g., novel animals or fictitious social groups) as training stimuli (Huijding et al., 2009; Van Dessel et al., 2018). Thus, it remains unclear whether a preexisting approach bias toward soft drink consumption is necessary for AAT effects.

Conclusions

Habitual consumption of sugary beverages is a major risk factor for the development of several chronic diseases. Finding new behavioral intervention methods to reduce soda consumption is therefore an important task for nutrition and health sciences. The present study examined whether explicit training of approach and avoidance responses to soda drinks will change the consumption of soft drinks immediately after the training in a randomly selected student sample. In four experiments, lemonade consumption and preference scores did not change after extended AAT training. Importantly, this result was obtained despite high total power (N = 300) and despite evidence that the AAT procedure was effective in producing a transient change in automatic action tendencies (Experiment 3). Furthermore, Bayesian analyses provided substantial evidence for a null model of AAT effects. While finding no training effect may come as a disappointing finding, we believe that this knowledge could be an important caution for nutrition educators and health practitioners seeking new intervention techniques.

References

Amir, N., Kuckertz, J. M., & Najmi, S. (2013). The effect of modifying automatic action tendencies on overt avoidance behaviors. *Emotion*, *13*(3), 478–484. https://doi.org/10.1037/a0030443

Block, J. P., Chandra, A., McManus, K. D., & Willett, W. C. (2010). Point-of-purchase price and education intervention to reduce consumption of sugary soft drinks. *American Journal of Public Health*, *100*(8), 1427–1433. https://doi.org/10.2105/AJPH.2009.175687

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2013). Approach bias modification in alcohol dependence: Do clinical effects replicate and for whom does it work best? *Developmental Cognitive Neuroscience*, *4*(Supplement C), 38–51. https://doi.org/10.1016/j.dcn.2012.11.002

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2014). Implementation of approach bias re-training in alcoholism—How many sessions are needed? *Alcoholism: Clinical and Experimental Research*, *38*(2), 587–594. https://doi.org/10.1111/acer.12281

Eder, A. B. (2011). Control of impulsive emotional behaviour through implementation intentions. *Cognition & Emotion*, *25*(3), 478–489. https://doi.org/10.1080/02699931.2010.527493

Eder, A. B., & Rothermund, K. (2008). When do motor behaviors (mis)match affective stimuli? An evaluative coding view of approach and avoidance reactions. *Journal of Experimental Psychology: General*, *137*(2), 262–281.

Eder, A. B., Rothermund, K., & Proctor, R. W. (2010). The prepared emotional reflex: Intentional preparation of automatic approach and avoidance tendencies as a means to regulate emotional responding. *Emotion*, *10*(4), 593–598.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229–238.

Field, M., Caren, R., Fernie, G., & De Houwer, J. (2011). Alcohol approach tendencies in heavy drinkers: Comparison of effects in a relevant stimulus-response compatibility task and an approach/avoidance Simon task. *Psychology of Addictive Behaviors*, *25*(4), 697–701. https://doi.org/10.1037/a0023285

Field, M., Kiernan, A., Eastwood, B., & Child, R. (2008). Rapid approach responses to alcohol cues in heavy drinkers. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*(3), 209–218. https://doi.org/10.1016/j.jbtep.2007.06.001

Fishbach, A., & Shah, J. Y. (2006). Self-control in action: Implicit dispositions toward goals and away from temptations. *Journal of Personality and Social Psychology*, *90*(5), 820–832.

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., … Pine, D. S. (2010). Attention bias modification treatment: A meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry*, *68*(11), 982–990. https://doi.org/10.1016/j.biopsych.2010.07.021

Hsee, C. K., Tu, Y., Lu, Z. Y., & Ruan, B. (2014). Approach aversion: Negative hedonic reactions toward approaching stimuli. *Journal of Personality and Social Psychology*, *106*(5), 699–712. https://doi.org/10.1037/a0036332

Huijding, J., Field, A. P., De Houwer, J., Vandenbosch, K., Rinck, M., & van Oeveren, M. (2009). A behavioral route to dysfunctional representations: The effects of training approach or avoidance tendencies towards novel animals in children. *Behaviour Research and Therapy*, *47*(6), 471–477. https://doi.org/10.1016/j.brat.2009.02.011

Janssen, T., Larsen, H., Vollebergh, W. A. M., & Wiers, R. W. (2015). Longitudinal relations between cognitive bias and adolescent alcohol use. *Addictive Behaviors*, *44*, 51–57. https://doi.org/10.1016/j.addbeh.2014.11.018

JASP Team. (2018). JASP (Version 0.8.5). Abgerufen von https://jasp-stats.org

Jeffreys, H. (1961). *Theory of probability*. Oxford, England: Oxford University Press.

Kakoschke, N., Kemps, E., & Tiggemann, M. (2017). 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

Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology*, *92*(6), 957.

Lindgren, K. P., Wiers, R. W., Teachman, B. A., Gasser, M. L., Westgate, E. C., Cousijn, J., … Neighbors, C. (2015). Attempted training of alcohol approach and drinking identity associations in US undergraduate drinkers: Null results from two Studies. *PLOS ONE*, *10*(8), e0134642. https://doi.org/10.1371/journal.pone.0134642

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*(4), 276.

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology*, *111*(1), 107–123. https://doi.org/10.1037/0021-843X.111.1.107

Mensink, G. B., Schienkiewitz, A., Rabenberg, M., Borrmann, A., Richter, A., & Haftenberger, M. (2018). Consumption of sugary soft drinks among children and adolescents in Germany. Results of the cross-sectional KiGGS Wave 2 study and trends. *Journal of Health Monitoring*, *3*(1), 31–36.

Mitchell, C., & Pelley, M. L. (2010). *Attention and associative learning: From brain to behaviour*. Oxford University Press UK.

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2014). Prevalence of childhood and adult obesity in the United States, 2011-2012. *JAMA*, *311*(8), 806–814. https://doi.org/10.1001/jama.2014.732

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532-555.

Peeters, M., Wiers, R. W., Monshouwer, K., van de Schoot, R., Janssen, T., & Vollebergh, W. A. M. (2012). Automatic processes in at-risk adolescents: The role of alcohol-approach tendencies and response inhibition in drinking behavior. *Addiction, 107*(11), 1939–1946. https://doi.org/10.1111/j.1360-0443.2012.03948.x

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 105–120.

Snelleman, M., Schoenmakers Tim, M., & Mheen, D. (2015). Attentional Bias and Approach/Avoidance Tendencies Do Not Predict Relapse or Time to Relapse in Alcohol Dependency. *Alcoholism: Clinical and Experimental Research*, *39*(9), 1734–1739. https://doi.org/10.1111/acer.12817

Schumacher, S. E., Kemps, E., & Tiggemann, M. (2016). Bias modification training can alter approach bias and chocolate consumption. *Appetite*, *96*(Supplement C), 219–224. https://doi.org/10.1016/j.appet.2015.09.014

Spruyt, A., De Houwer, J., Tibboel, H., Verschuere, B., Crombez, G., Verbanck, P., … Noël, X. (2013). On the predictive validity of automatically activated approach/avoidance tendencies in abstaining alcohol-dependent patients. *Drug and Alcohol Dependence, 127*(1–3), 81–86. https://doi.org/10.1016/j.drugalcdep.2012.06.019

Stritzke, W. G., McEvoy, P. M., Wheat, L. R., Dyer, K. R., & French, D. J. (2007). The yin and yang of indulgence and restraint: The ambivalence model of craving. In P. W. O’Neal (Ed.), *Motivation of health behavior* (pp. 31–47). Nova Science Publishers.

Townshend, J., & Duka, T. (2001). Attentional bias associated with alcohol cues: Differences between heavy and occasional social drinkers. *Psychopharmacology*, *157*(1), 67–74. https://doi.org/10.1007/s002130100764

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison Wesley.

Van Dessel, P., De Houwer, J., & Gast, A. (2016). Approach–avoidance training effects are moderated by awareness of stimulus–action contingencies. *Personality and Social Psychology Bulletin*, *42*(1), 81–93. https://doi.org/10.1177/0146167215615335

Van Dessel, P., De Houwer, J., Gast, A., & Tucker Smith, C. (2014). Instruction-based approach-avoidance effects: Changing stimulus evaluation via the mere instruction to approach or avoid stimuli. *Experimental Psychology*, *62*(3), 161–169. https://doi.org/10.1027/1618-3169/a000282

Van Dessel, P., De Houwer, J., Roets, A., & Gast, A. (2016). Failures to change stimulus evaluations by means of subliminal approach and avoidance training. *Journal of Personality and Social Psychology, 110*(1), e1–e15. https://doi.org/10.1037/pspa0000039

Van Dessel, P., Eder, A., & Hughes, S. J. (2018). Mechanisms underlying effects of approach-avoidance training on stimulus evaluation. *Journal of Experimental Psychology: Learning, Memory and Cognition*.

van Hemel-Ruiter, M. E., de Jong, P. J., & Wiers, R. W. (2011). Appetitive and regulatory processes in young adolescent drinkers. *Addictive Behaviors*, *36*(1), 18–26. https://doi.org/10.1016/j.addbeh.2010.08.002

Vartanian, L. R., Schwartz, M. B., & Brownell, K. D. (2007). Effects of soft drink consumption on nutrition and health: A systematic review and meta-analysis. *American Journal of Public Health, 97*(4), 667–675. https://doi.org/10.2105/AJPH.2005.083782

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. *Psychological Science*, *22*(4), 490–497. https://doi.org/10.1177/0956797611400615

Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., & Ridderinkhof, K. R. (2013). Cognitive bias modification and cognitive control training in addiction and related psychopathology: Mechanisms, clinical perspectives, and ways forward. *Clinical Psychological Science*, *1*(2), 192–212. https://doi.org/10.1177/2167702612466547

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

Wiers, R. W., Houben, K., Fadardi, J. S., van Beek, P., Rhemtulla, M., & Cox, W. M. (2015). Alcohol cognitive bias modification training for problem drinkers over the web. *Addictive Behaviors*, *40*, 21–26. https://doi.org/10.1016/j.addbeh.2014.08.010

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.

Zogmaister, C., Perugini, M., & Richetin, J. (2016). Motivation modulates the effect of approach on implicit preferences. *Cognition and Emotion*, *30*(5), 890–911. https://doi.org/10.1080/02699931.2015.1032892