

# The Cue-Approach Task as a General Mechanism for Long-Term Non-Reinforced Behavioral Change

Tom Salomon<sup>1</sup>, Rotem Botvinik-Nezer<sup>1,2</sup>, Tony Gutentag<sup>3</sup>, Rani Gera<sup>1</sup>, Roni Iwanir<sup>1</sup>, Maya Tamir<sup>3</sup>, Tom Schonberg<sup>\*1,2</sup>

Tel Aviv University

## Author Notes

This work was supported by the Israeli Science Foundation (ISF number 1798/15) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 715016) granted to Tom Schonberg.

\*Correspondence concerning this article should be addressed to Tom Schonberg, Department of Neurobiology, Faculty of Life Sciences, and Sagol School of Neuroscience, Tel Aviv University, Ramat Aviv 6139001, Tel Aviv, Israel. Email: Schonberg@post.tau.ac.il.

## Author Contributions

T.Sa. designed experiments, collected and analyzed data and wrote the manuscript. R.B.N. contributed to the design of the experiments and write-up of manuscript. T.G, R.G and R.I. collected data and assisted with manuscript write-up. M.T. contributed to the design of the experiments and manuscript write-up; T.Sc. design experiments, assisted with data analysis and wrote the manuscript with T.Sa. All writers actively contributed to the writing process of the manuscript.

<sup>1</sup>Department of Neurobiology, Tel Aviv University. <sup>2</sup>Sagol School of Neuroscience, Tel Aviv University. <sup>3</sup>Department of Psychology, Hebrew University of Jerusalem

\*Corresponding author

### **Abstract**

Recent findings show that preferences for food items can be modified without external-reinforcements using the cue-approach task. In the task, the mere association of food item images with a neutral auditory cue and a speeded button press, resulted in enhanced preferences for the associated stimuli. Here, in a series of 10 independent samples with a total of 255 participants, we show we can enhance preferences using this non-reinforced method for faces, fractals and affective images as well as snack foods, using auditory, visual and even aversive cues. This change was highly durable in follow-up sessions performed one to six months after training. Preferences were successfully enhanced for all conditions, except for negative valence items. These findings promote our understanding of non-reinforced change, suggest a boundary condition for the effect and lay the foundation for development of novel applications.

Behavioral change is an essential tool to improve health and quality of life, from treating addictions to eating and mood disorders<sup>1-3</sup>. Scientific research on behavioral change mainly focused on the effects of external reinforcements<sup>2,4-7</sup> or altering the presentation of the decision problem<sup>8-10</sup>. Recently, the Cue-Approach Task (CAT)<sup>11</sup> was introduced as a successful method for enhancing preferences for food items, without external reinforcement, context change or self-control. Here, we test multiple hypotheses that are aimed to shed light on this mechanism by studying its generalizability to multiple stimuli and cues, as well as the long-term durability of the effect.

In the CAT, participants initially indicated their preferences for a set of snack-food items by specifying their willingness to pay for each item in an auction procedure. Then, in the cue-approach training, some of the items were consistently associated with a neutral auditory cue and a speeded button press response (these stimuli were termed ‘Go items’), whereas other stimuli were presented without a cue (‘No-Go items’). In the following probe phase, participants were asked to choose a snack they would like to eat at the end of the experiment. Each probe-choice comprised of a pair of items with similar initial values, in which one of the two snacks was a Go item. Results showed that the mere association of snack-food images with a neutral auditory cue and a speeded button press, resulted in enhanced preferences for Go items over No-Go items. This preference change effect varied across different value categories – resulting in enhanced preferences for snack-food items of initial high-value, yet significantly less prominent change in preferences for low-value items. The effect was maintained two-months following training<sup>11</sup>. Additional studies with CAT<sup>12</sup> found that for the behavioral change to take place, cue-approach training required the presence of both a speeded button press response and a cue; i.e. CAT had no effect when training was conducted with an early cue onset which was followed with a full one second to respond. In addition, when training included only button press responses with no cues, enhanced preferences for Go items were not observed<sup>12</sup>. It was also found that the CAT effect

induces preference changes beyond the hand motor circuit itself, relying on a study that trained participants with the button press, whereas choices during probe were made using the eye gaze<sup>12</sup>. Veling et al.<sup>13</sup> extended the range of consumable items for non-reinforced change to fruits and vegetables, and also found that the effect requires a time-restricted response. Finally, it was also shown that using CAT, it is possible to increase the odds of participants choosing low-value snack-food items over high-value items, in comparison to a baseline rate of choices between two non-cued items<sup>14</sup>.

Based on these findings, summarizing over 15 samples, the CAT has been established as a replicable paradigm to induce preference changes for consumable food items, without external reinforcement. The underlying mechanism of this preference change has not yet been fully elucidated. Neural findings showed a neural signature indicative of value change for the Go items in the ventromedial prefrontal cortex during choices<sup>11</sup>. Further imaging classification studies were not able to clearly define its underlying basic cognitive constructs, but did show the involvement of frontal control networks following training<sup>15</sup>. Currently it is suggested that the underlying mechanism relies on attention<sup>11,12,15</sup>, inspired by its conceptual resemblance to the attentional boost effect<sup>16</sup>, where memory for task-irrelevant stimuli was enhanced when stimuli co-appeared with an irrelevant target cue. Thus far, the CAT effect, was demonstrated only on consumable items with a neutral auditory cue and up to a two-months follow up. The goal of the present study was to test multiple hypotheses in 10 independent samples using various non-consumable stimuli and different cues, as well as long-term follow-up sessions. This would promote our understanding of the underlying mechanism, as well as the boundary conditions of the CAT as a general mechanism for long term non-externally reinforced change.

Our first research aim was to test whether CAT can enhance preferences for non-consumable items. To this end, we performed CAT with a variety of stimuli. First, we tested the efficacy of CAT in changing preferences for face images (Experiment 1), which are

important social stimuli and known to elicit preferences<sup>17,18</sup>. Second, we tested if CAT could be used to change preferences for more abstract stimuli, such as fractal art images (Experiment 2), which participants are unaccustomed to and are also more difficult to associate with semantic knowledge. Third, we continued to examine the importance of the stimuli's affective valence. Research with external reinforcement procedures<sup>19-21</sup> indicated an important interplay between the stimulus affective valence (aversive versus appetitive) and the required response (approach versus avoidance) – showing better association of approach responses to appetitive stimuli, and avoidance responses to aversive stimuli. To examine the importance of the stimuli's valence, we performed two experiments using affective images from the International Affective Picture System<sup>22</sup> (IAPS): one experiment was performed with positive IAPS stimuli (Experiment 3), and a second with negative IAPS stimuli (Experiment 4).

After examining the importance of different stimuli features (i.e., consumability, abstractness and affective valence), we sought to examine the importance of the cues' nature. Previous studies suggested that cues of different modalities may alter performance in different tasks<sup>23,24</sup>. Since all of the previous experiments with the CAT<sup>11-15</sup> were conducted with a neutral auditory cue, we examined whether the cue's auditory modality is a crucial feature of the CAT. In Experiment 5 we tested whether training with a visual cue would also induce enhanced preferences for familiar snack-food items. In Experiment 6, we further tested the importance of the valence<sup>19-21</sup> of the cue by associating appetitive snack-food items with an aversive tone cue. Here we did not have a clear hypothesis if the aversive cue combined with the appetitive stimuli will lead to enhanced preferences, as we put to the test two competing hypotheses – on the one hand, CAT with neutral cue resulted in enhanced preferences for the associated items, and therefore if the valence of the cue is not a cardinal feature, we would expect enhanced preferences also with an aversive cue. On the other hand,

the cue's valence may be a fundamental factor, as association of stimuli with an aversive cue may result in reduced preferences via classical conditioning<sup>25</sup>.

To adhere to principles of replicability<sup>26-28</sup> in Experiments 7-10 we performed improved replications of our novel findings with non-consumable items from Experiments 1-4. In all of the replications, we used a more extensive training protocol with both an auditory cue (Experiments 7-8) and a visual cue (Experiments 9-10). In Experiment 7 we replicated Experiment 1 with a new set of faces adapted from a more recent, better quality dataset. This was done following reports by participants that the stimuli used in Experiment 1 were outdated and not visually appealing. In Experiment 8 we replicated Experiment 2 with fractals. Finally, in Experiments 9-10 we replicated Experiments 3-4 with the positive and negative affective IAPS stimuli, with a visual cue during training. Replicating the results with a visual cue was important to ensure participants in the negative IAPS condition did not avert their gaze from the unpleasant images during training.

In addition to testing the generalizability of the effect with multiple non-consumable stimuli and cues, another central aim of the current study was to examine the long-term durability of non-reinforced behavioral change. Therefore, we invited participants from five experiments to an additional follow-up session. We assessed the long-term effect following a period of one-month (Experiment 7), two-months (Experiments 5 and 6), four-months (Experiment 2) and six-months (Experiment 8) after training. Demonstrating that the behavioral change effect is durable over a long period of time could point to its underlying mechanism that putatively effects the low-level representation of the items<sup>29</sup>, as well as potential applicability in inducing long-term change.

Based on previous findings<sup>11-15</sup>, we hypothesized that CAT would result in increased preferences for Go items (i.e. items that were previously associated with a cue and a response) over No-Go items with similar initial preferences. We also predicted that this behavioral change would be more robust for high-value items (items for which a participant had a higher

pre-existing preference) than for low-value items, as reported in most studies with CAT<sup>11,12,15</sup>, though not all of them<sup>13</sup>.

## Methods

### Participants

A total of 255 healthy participants participated in one of 10 independent experiments. In five experiments (Experiments 2 and 5-8), 84 participants agreed to return for an additional follow-up session (average of 67% retention rate of the first samples), one to six-months following their original participation date (for a demographic description of each experimental sample see Table 1). All participants had normal or corrected to normal vision and hearing, and gave their informed consent to participate in the experiments in return for monetary compensation or in return for course credit (course credit was granted to some of the participants in Experiments 3-4 and 9-10). The study was approved by the ethics committee of Tel Aviv University (Experiments 1-2 and 5-8) and by the ethics committee of the Hebrew University of Jerusalem (Experiments 3-4, and 9-10).

Our target sample size of  $n=25$  for Experiments 2-10 was selected based on power analysis of 80% power to detect an effect with .05 significance. Power was calculated using a two-sided one-sample t-test, comparing the proportion of trials high-value Go items were chosen at 50% chance. Data were obtained from previous experiments with the CAT<sup>11</sup> and Experiment 1 (power analysis was conducted using R's pwr package<sup>30</sup>, and is available online at [osf.io/h36vr](https://osf.io/h36vr)). Final sample sizes ranged from 25-29; Data from participants beyond  $n=25$  were collected in order to guarantee the minimal  $n=25$  in case of exclusion following primary quality assurance of the data. All reported effects remain consistent when the participants beyond  $n=25$  were removed.

Table 1

*Demographics*

Experiment Stimulus	Go cue	Sample Size (Excluded)	Females (Proportion)	Mean Age (SD)	Follow-up	
					Sample Size	Mean Interval <sup>a</sup> (Range)
Exp. 1: Faces	Auditory	19 (12)	10 (53%)	24.42 (3.27)	-	-
Exp. 2: Fractals	Auditory	25 (3)	15 (60%)	23.04 (1.92)	15	116.2 (73-170)
Exp. 3: Positive IAPS	Auditory	27 (1)	17 (63%)	24.11 (2.62)	-	-
Exp. 4: Negative IAPS	Auditory	28 (5)	21 (75%)	25.04 (2.53)	-	-
Exp. 5: Snacks	Visual	25 (1)	16 (64%)	22.24 (2.30)	21	51.8 (34-69)
Exp. 6: Snacks	Auditory Aversive	25 (5)	15 (60%)	24.40 (2.15)	14	61.5 (31-105)
Exp. 7: Faces	Auditory	25 (1)	18 (72%)	25.16 (2.44)	18	33.2 (22-56)
Exp. 8: Fractals	Auditory	25 (2)	19 (76%)	24.20 (3.33)	16	183.9 (171-203)
Exp. 9: Positive IAPS	Visual	29 (3)	18 (62%)	23.27 (2.08)	-	-
Exp. 10: Negative IAPS	Visual	27 (3)	20 (74%)	23.67 (2.60)	-	-

<sup>a</sup> Mean interval in days from training to follow-up session.

In addition to the reported 255 participants, across the 10 experiments 31 additional participants were excluded from final analyses (see Table 1 and Supplementary Table 1). Twenty-one participants were disqualified due to poor performance in training, an exclusion criteria adopted from previous CAT studies<sup>11,12</sup>; six participants due to technical problems with the apparatus running the experiment; two participants requested to quit; one participant entirely avoided choices of low-value snacks during the probe phase and one participant due



to extreme intransitivity in initial preferences (transitivity score from initial preferences evaluation of 3.67 *SD* below the group mean).

## Materials

**Stimuli.** Six different stimuli sets were used, each containing 60 identically-sized color images of either unfamiliar faces<sup>31,32</sup> (two different datasets were used in Experiment 1 and Experiment 7), fractal art<sup>33</sup> (Experiments 2 and 8), images from the International Affective Picture System (IAPS)<sup>22</sup> of positive valence (Experiments 3 and 9) or negative valence (Experiments 4 and 10) and popular Israeli snack-food items<sup>34</sup> (dataset created in our laboratory for Experiments 5 and 6).

For Experiment 1 we used face images adapted from a functional MRI face localizer task<sup>31</sup>. The face stimuli included images of 30 male and 30 female front-facing individuals on a white background, sized  $280 \times 296$  pixels. In Experiment 7, we performed an improved replication of the face experiment using a newer dataset<sup>32</sup> with better quality face images. Stimuli included 30 male and 30 female front-facing figures, posing a neutral expression with limited facial hair and make-up. The original images were graphically edited using Adobe Photoshop, cropped to identical size ( $400 \times 500$  pixels) and the original green screen background was replaced with a homogenous gray background. In order to maintain similar position and relative proportion of the stimuli, faces were centered and aligned according to the location of the pupils (see example in Figure 1a). Computer-generated fractal art stimuli<sup>33</sup> used in Experiments 2 and 8 included 60 identical sized ( $576 \times 432$  pixels) color images.



*Figure 1.* Stimuli and cue examples. 1a. Example of a face stimulus with neutral expression used in Experiment 7. 1b. Example of snack-food stimulus used in Experiments 5-6. 1c. Semi-transparent Gabor used as a visual cue in Experiments 6 and 9-10. 1d. Example of snack-food stimulus with the visual cue overlaid on it as presented

For Experiments 3-4 and 9-10 with IAPS stimuli, images were selected based on the norms published by Lang et al.<sup>22</sup>. For the positive IAPS experiments (Experiments 3 and 9) we used affective images, which were rated as inducing positive affect (valence  $M = 7.10$  on a 1 to 9 scale,  $SD = 0.43$ ; arousal  $M = 5.16$  on a 1 to 9 scale,  $SD = 0.46$ ). For the negative IAPS experiments (Experiments 4 and 10) we used images rated as inducing negative affect (valence  $M = 2.87$ ,  $SD = 0.55$ ; arousal  $M = 5.50$ ,  $SD = 0.78$ ). Arousal ratings were similar in both datasets. In experiments with negative affective stimuli we also used three additional stimuli not from the official IAPS dataset, which were found to induce negative affect in another independent study<sup>35</sup>. All IAPS stimuli were rescaled to identical size ( $533 \times 400$  pixels).

For Experiments 5 and 6, conducted with familiar snack-food items, we prepared a new stimuli dataset of local familiar snacks<sup>34</sup>. Snack-food images were of identical size ( $576 \times 432$  pixels) and presented on homogenous black background.

**Cues.** We used three types of training cues: In Experiments 1-4, and 7-8, we used the original neutral auditory cue of a 180-ms sinusoidal wave tone, identical to the one used in previous CAT studies<sup>11</sup>; in Experiments 5, and 9-10, we used a visual cue of a 180-ms semi-transparent Gabor (see example in Figure 1c-d); in Experiment 6 we used an aversive 300-ms auditory cue created with a cotangent function (provided as Supplementary Code).

**Stimuli presentation.** Stimuli were presented using MATLAB (Mathworks, Inc. Natick, MA, USA), Psychtoolbox<sup>36</sup> and Python-based Pygame<sup>37</sup> packages, on 21.5 inch iMac or PC with a 21 inch screen. In order to induce the aversive cue in Experiment 6, the sound was played in a controlled high volume, using Plantronics BackBeat Pro noise-canceling headphones.

## Procedure

**Baseline evaluation of subjective preferences.** Participants' baseline subjective preferences for each of the stimuli in a given experiment were evaluated individually using

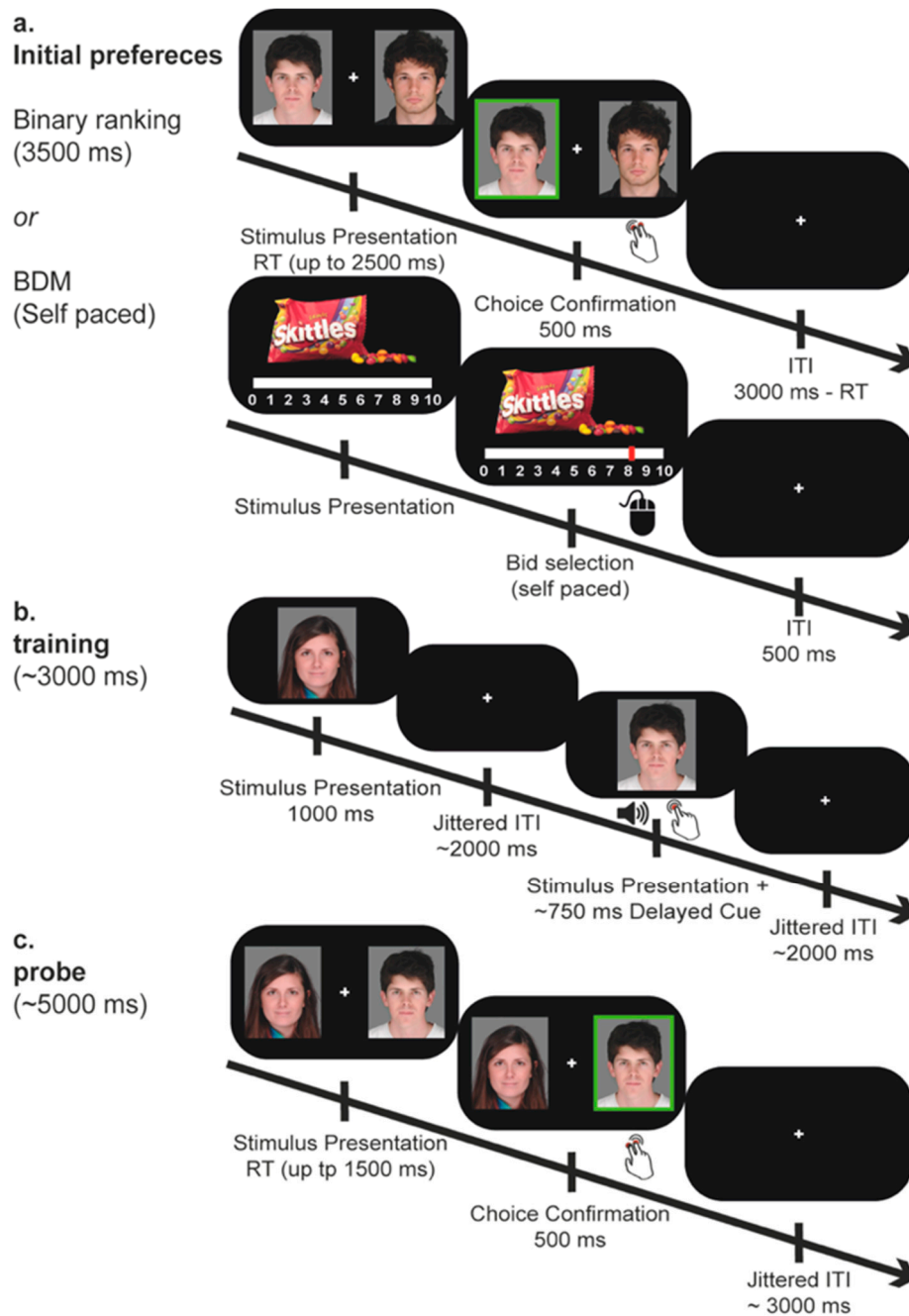
two methods: an auction procedure in experiments with consumable snack-food items (Experiments 5 and 6), and a forced-choice binary ranking procedure for non-consumable items (Experiments 1-4 and 7-10). Based on these valuations, stimuli were rank-ordered individually for each participant, from the most liked item (rank 1) to the least liked item (rank 60).

*Auction procedure for snack-food items.* In experiments with snack-food items (Experiments 5 and 6), participants underwent an auction task based on the Becker DeGroot Marschak (BDM)<sup>38</sup> auction procedure, to obtain participants' willingness to pay, similarly to previous CAT studies<sup>11-15</sup>. Prior to their participation in the experiment, subjects were asked to fast for at least four hours. Upon arrival to the laboratory, participants received a sum of 10 New Israeli Shekels (~2.7\$ US equivalent) to be used in the auction task. In this task, snack-food items were presented individually on the screen one at a time. Participants selected their bid for each item on a visual analog scale using a mouse (see Figure 2a). Participants were explicitly told that offering the amount they are willing to pay for each item was the best strategy for the task. Participants were informed in advance, that at the end of the experiment one trial will be chosen at random to be played out for real purchase from the experimenter. The computer randomly generated a counter bid, ranging from 0-10 with 0.5 increments. In case the participant's bid was higher than the computer's, he or she had won the bid and were required to buy the item in return for the computer bid price; otherwise, the participant could not purchase the item, but was left with the money designated for the auction at the beginning of the experiment.

*Binary ranking for non-consumables items.* In experiments with non-consumable stimuli (i.e. faces, fractals and IAPS; Experiments 1-4 and 7-10), we used a forced-choice binary ranking procedure. In this task, 60 stimuli were randomly paired with each other to form 300 unique pairs. For each pair of stimuli, participants had 2500-ms to choose their preferred stimulus, followed by a 500-ms choice confirmation and 500-ms fixation cross (see

Figure 2a). In order to maintain balanced exposure, each stimulus was presented in exactly 10 pairs during the binary ranking phase.

Choices were then quantified into ranking scores. Based on the assumption of choice transitivity from the Rational Choice Theory<sup>39</sup>, we used the outcomes from the set of binary choices in order to deduce individual preferences for the presented set of stimuli. That is, if stimulus A is preferred over B and stimulus B is preferred over C, then their respective ranks follows  $A > B > C$ . In order to maximize ranking validity and specificity we used the Colley Matrix algorithm<sup>40</sup>, designed to solve ranking problems with a relatively small number of binary outcomes. This binary ranking procedure resulted in a ranking list of the 60 stimuli, based on each participant's individual preferences.



*Figure 2.* An example of the experimental procedure outline using face stimuli. Mean trial duration written in parenthesis, tilde sign indicates varying duration. 2a. Initial preferences evaluation using either a binary ranking procedure (for non-consumable stimuli) or an auction procedure (for snack-food items, in Experiments 5-6). 2b. In the training phase, Go items were consistently paired with a cue and a speeded button press. 2c. In the probe phase participants chose their preferred stimulus between pairs of items with similar initial value, where only one was a Go item, previously associated with a cue during training.

Colley Matrix ranking scores typically range in scores from 0 (least liked) to 1 (most liked), with a fixed mean of 0.5. An intransitive choice pattern is characterized by densely distributed scores around the center of 0.5, while a distinct preferences pattern leads to more distributed ranking scores. From these rankings, we quantified a transitivity score for each participant as the standard deviation of the participant's ranking scores. Participants who demonstrated extreme intransitive choice patterns (3 *SD* below the group mean), were excluded from final analyses.

**Cue-approach training.** Following the baseline evaluation of subjective preferences, participants underwent a 40-minutes long cue-approach training procedure, during which a consistent association was formed between some of the stimuli and a Go cue (see Figure 2b). Each stimulus in the training set was presented individually on the screen for 1000-ms, once during each training run. Stimuli were randomly ordered and followed by a jittered fixation cross with an average duration of 2000-ms (*SD* = 1243-ms; range of 1000 - 6000-ms, 1000-ms intervals).

We used two formats of training designs, both were 40-minutes long, but the number of training runs (i.e. total number of presentations of each stimulus) was different, as well as the number of cued Go items. In the first, shorter-training design, used in Experiments 1-4 the training set consisted of all 60 stimuli (as in the original CAT study<sup>11</sup>), presented in 12 training runs. In the second, more extensive training design, used in Experiments 5-10, the number of training runs was increased to 20 and the duration of each run was reduced by presenting only a subset of 40 stimuli, consisting of 20 high-value (ranked 3-22, above the median rank) and 20 low-value items (ranked 39-58).

One third of the training set items (20 out of 60 items in the shorter-training design) or 30% of items (12 out of 40 items in the extensive-training design) were associated with the Go cue. Participants were instructed to respond to the Go cue by pressing a keyboard button as fast as possible, before stimulus offset. Participants were not informed in advance

that the association of stimuli with the cue would be consistent or which items would be Go items.

Items were assigned to be associated with the Go cue based on the previous baseline preferences evaluation task. Two sets of high-value stimuli with identical mean ranks and two sets of low-value stimuli with identical mean ranks were predetermined (see Supplementary Figure 1). For each participant, one high-value set and one low-value set were chosen to be consistently associated with the cue (Go items), whereas all other stimuli appeared without a cue (No-Go items). The cue appeared following a delay of ~750-ms from stimulus onset. To maintain a balanced difficulty level throughout the training phase, the delay was modified according to participants' performances, as conducted in previous CAT studies<sup>11</sup>.

**Probe.** Preference change following CAT was evaluated in a probe phase. On each probe trial, two items appeared to the right and left of a central fixation cross and participants were asked to select their preferred stimulus. In each pair, both items were of similar initial value (either high-value or low-value), but only one item was a Go item, associated with a cue during training. For each pair, participants had 1500-ms to select their preferred stimulus, followed by a 500-ms choice confirmation and a fixation cross for a jittered duration with an average of 3000-ms (range of 1000-11000-ms, 1000-ms intervals; see Figure 2c).

In the design used in Experiments 1-4, probe choices of each value category (high and low value) included eight Go items which were compared to eight No-Go items of equal mean rank, for a total of 64 (8x8) unique comparisons per value category. In the second design used in Experiments 5-10, probe choices included six Go and six No-Go items for a total of 36 (6x6) unique probe comparisons in each category (Supplementary Figure 1). While previous studies with CAT<sup>11-15</sup> only included 4x4 unique comparisons per value category, in Experiments 1-4 of the current work we modified the design of the training and probe parts to include 8X8 Go items, in an attempt to increase the number of unique probe choices.

Following experiments 1-4, we hypothesized that a value category of 16 items was too wide, as within each value-category, participants were more likely to choose the relatively higher valued items. Therefore, in the later designs (used in Experiments 5-10) the number of Go items in each value-category was reduced to 6, in order to maintain smaller gaps in initial preferences between all items in the same value category.

In addition to these comparisons, as in previous CAT experiments<sup>11-15</sup>, ‘sanity check’ trials were also incorporated in the probe phase. In the ‘sanity check’ trials, participants were asked to choose between pairs of items in which one item was of initial high-value and the other of initial low-value (both Go or both No-Go items), in order to reassure the validity of the initial preference evaluation procedure. The probe phase included two runs. On each run, all unique probe pairs were presented in a random order.

In Experiments 5-6, conducted with snack-food items, choices were made for actual food consumption. To ensure incentive-compatible choices, participants were informed in advance that one trial would be selected at random, and that they would receive the item selected on that trial at the end of the experiment.

**Memory.** At the end of the experiment, participants performed an Old/New and Go/No-Go recognition tasks. The results of these tasks are not reported here as they are beyond the scope of this paper.

**Follow-up.** In Experiments 2 and 5-8, participants returned after a predetermined mean period of either one, two, four or six months (see Table 1). In the follow-up sessions, participants completed only the probe and memory tasks again. For each participant, the follow-up tasks included the same stimuli and probe pairs he or she had previously performed during the original session, presented in a random order. All participants were notified in advance and encouraged to return to the follow-up sessions.



## Results

To assess preference changes following training, we analyzed the proportion of probe phase trials in which participants preferred the Go items over the No-Go items, using a two-tailed repeated measures logistic regression. In each pair, both items were of similar initial preference based on the baseline evaluation phase. We hypothesized that the cue approach effect would enhance preferences for the Go items above the chance level of 50% of trials (log-odds = 0; odds-ratio = 1). For each logistic regression analysis, the estimated odds ratios (OR) are reported as the corresponding effect sizes along with 95% confidence intervals (CI). Analyses were conducted separately for pairs of high-value items and pairs of low-value items, similar to previous CAT studies<sup>11–15</sup>. The results of all experiments are summarized in Table 2, and Figures 3-4. Analyses and visualizations were conducted using lme4<sup>41</sup> and ggplot2<sup>42</sup> R packages, and are available online along with the experimental data at [osf.io/puxhx](https://osf.io/puxhx).

### Experiments 1-4: Changing Preferences for Non-Consumable Items - Faces, Fractals and Affective Stimuli

The first goal of the current work was to examine CAT's efficacy in changing preferences for non-consumable stimuli. To achieve this goal, we tested CAT with unfamiliar stimuli of faces in Experiment 1, as well as abstract stimuli of fractal art in Experiment 2. Experiments 3-4 aimed to test if the CAT effect is dependent on the stimuli's valence by testing the ability to change preferences for positive (Experiment 3) and negative (Experiment 4) affective stimuli.

In Experiment 1, Cue approach training with unfamiliar faces resulted in no significant preferences for high-value Go items over high-value No-Go items (mean proportion = 52.4%, OR = 1.11, 95% CI [0.89, 1.38],  $p = .343$ ); and a small preference of low-value Go items over low-value No-Go items (mean proportion = 57.1%, OR = 1.39, 95% CI [1.05, 1.84],  $p = .023$ ; see Figure 3 and Table 2). A significant difference was found

between high and low value probe choices with a greater effect for the low-value items (OR = 0.82, 95% CI [0.73, 0.93],  $p = 0.001$ ).

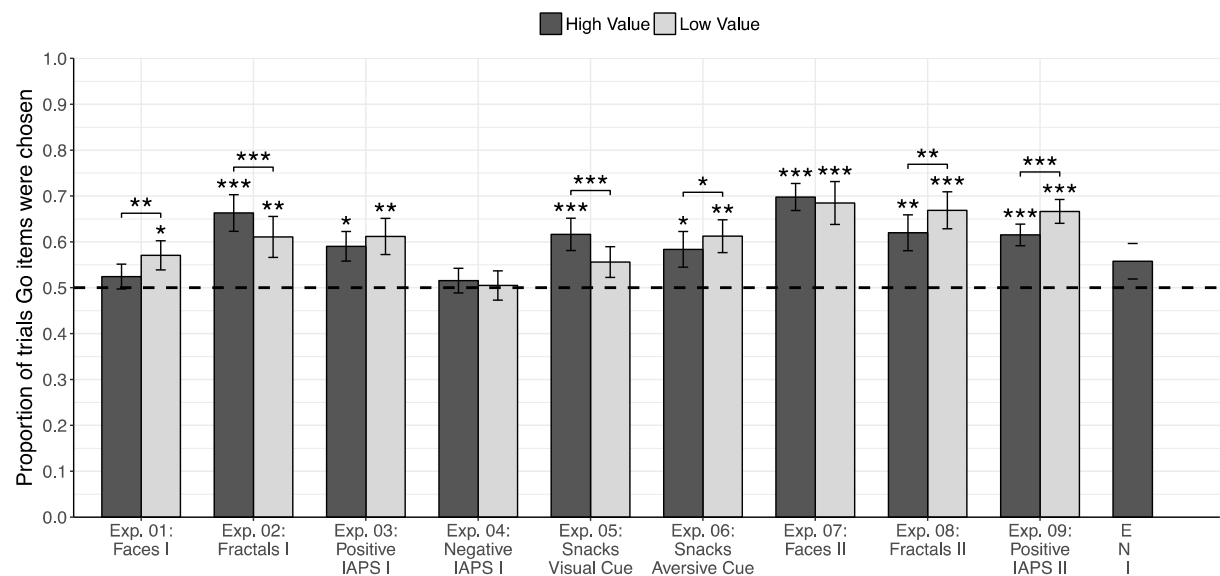


Figure 3. Probe results of Experiments 1-10.

Mean proportion of trials participants chose Go items over No-Go items, for high-value (dark gray) and low-value (light gray) probe pairs. Dashed line indicates 50% chance level, error bars represent standard error of the mean. Asterisks reflect statistical significance in a two-tailed logistic regression analysis. Asterisks above each bar represent proportions different from chance (log-odds = 0, odds-ratio = 1). Asterisks above pairs of bars represent differential effect for the two value categories (log-odds = 0, odds-ratio =

Table 2

*Probe Results - Mean Proportion of Trials Participants Chose Go Items Over No-Go Items*

Experiment	Go cue	Training Runs	Mean Proportion Go items were Chosen			
			First session		Follow-up Session	
			High-Value	Low-Value	High-Value	Low-Value
Exp. 1: Faces	Auditory	12	52.4%, $p = .343$	57.1%, $p = .023$	-	-
Exp. 2: Fractals	Auditory	12	66.3%, $p = 2.4E^{-4}$	61.1%, $p = .009$	58.1%, $p = .029$	54.5%, $p = .213$
Exp. 3: Positive IAPS	Auditory	12	59%, $p = .011$	61.2%, $p = .007$	-	-
Exp. 4: Negative IAPS	Auditory	12	51.6%, $p = .582$	50.5%, $p = .877$	-	-
Exp. 5: Snacks	Visual	20	61.7%, $p = .001$	55.6%, $p = .096$	56.9%, $p = .038$	53.2%, $p = .317$
Exp. 6: Snacks	Auditory Aversive	20	58.4%, $p = .030$	61.2%, $p = .002$	59.4%, $p = .031$	59.0%, $p = .066$
Exp. 7: Faces	Auditory	20	69.8%, $p = 8.8E^{-9}$	68.5%, $p = 2.3E^{-4}$	66.8%, $p = 4.8E^{-5}$	62.1%, $p = .027$
Exp. 8: Fractals	Auditory	20	62%, $p = .003$	66.9%, $p = 8.5E^{-5}$	60.8%, $p = .03$	60.8%, $p = .032$
Exp. 9: Positive IAPS	Visual	20	61.5%, $p = 1.4E^{-6}$	66.7%, $p = 1.2E^{-9}$	-	-
Exp. 10: Negative IAPS	Visual	20	55.8%, $p = .116$	53.3%, $p = .322$	-	-

*Note.* Reported  $p$  values indicate a significant deviation from chance level (mean proportion = 50% odds ratio = 1) in a two tailed repeated measures logistic regression

One of the important differences of this Experiment 1 from previous CAT studies<sup>11,12</sup> with appetitive snack-food stimuli was the increase in the range of each value category to 16 items (eight Go and eight No-Go items), in an attempt to increase the number of possible comparisons obtain better measurements. This increase came with a known cost of introducing choices between pairs of less homogenous initial values. As this experiment was different in design, we additionally performed an analysis to examine the results using only

the most extreme stimuli in each category, i.e. the highest four Go and four No-Go stimuli in the high-value category, and the lowest four Go and four No-Go stimuli in the low-value category. These analyses are equivalent to probe-choices in previous studies<sup>11-13</sup>. Examining the extreme choices revealed enhanced preferences for the highest Go items over the highest No-Go items (mean proportion = 57.9%, OR = 1.41, 95% CI [1.04, 1.92],  $p = .026$ ), and a similar trend for low-value choices (mean proportion = 57.9%, OR = 1.44, 95% CI [0.97, 2.17],  $p = .073$ ).

The results of Experiments 2-4 were also analyzed separately for the highest and lowest probe-choices as has been done in previous studies<sup>11-13</sup>. In these experiments, the additional analyses of the more extreme comparisons showed consistent results with the main analyses for the entire value category. Therefore, for these samples we do not report the additional analysis and they are available online ([osf.io/h36vr](https://osf.io/h36vr)), along with the data and all other analyses.

In Experiment 2, following cue-approach training with fractal art stimuli, participants consistently preferred the high-value Go items over the high-value No-Go items during probe phase (mean proportion = 66.3%, OR = 2.56, 95% CI [1.55, 4.22],  $p = 2.4E^{-4}$ ); similarly, participants consistently preferred the low-value Go items over the low-value No-Go items (mean proportion = 61.1%, OR = 1.92, 95% CI [1.18, 3.14],  $p = .009$ ). CAT effect on preferences was significantly stronger for high-value probe choices over low-value probe choices (OR = 1.31, 95% CI [1.17, 1.46],  $p = 2.2E^{-6}$ ).

In Experiment 3, following CAT with positive IAPS, participants significantly preferred Go over No-Go items, both in the high-value (mean proportion = 59.0%, OR = 1.62, 95% CI [1.12, 2.34],  $p = .011$ ) and in the low-value probe choices (mean proportion = 61.2%, OR = 2.06, 95% CI [1.22, 3.46],  $p = .007$ ). Differences between high and low value probe choices were trending towards more robust effects for low-value items (OR = -0.1, 95% CI [-0.20, 0.01],  $p = 0.067$ ).

In Experiment 4, conducted with negative IAPS stimuli, participants displayed no preferences for the Go items, neither for the high-value pairs (mean proportion = 51.6%, OR = 1.07, 95% CI [0.85, 1.34],  $p = .582$ ); nor the low-value pairs (mean proportion = 50.5%, OR = 1.02, 95% CI [0.77, 1.36],  $p = .877$ ). No significant difference between high and low value probe choices was found (OR = 1.05, 95% CI [0.95, 1.16],  $p = 0.316$ ).

### **Experiments 5-6: Using Cues of Different Modality and Valence**

After examining CAT with different stimuli, we went on to examine whether using an auditory cue was a requirement to induce CAT preference change. Therefore, in Experiment 5, we used a semi-transparent Gabor on top of snack-food items (See Figure 1c and 1d). In Experiment 6, following the results of Experiment 4 showing that CAT did not alter preferences for negative stimuli, we tested the importance of cue neutrality, whether the CAT could enhance preferences for appetitive snacks, even when cued with an aversive tone. In both Experiments 5 and 6, we used familiar local snack-food items, similar to those used in previous CAT studies<sup>11-15</sup>.

In Experiment 5, following training with a neutral visual cue, participants preferred the high-value Go items over the No-Go items (mean proportion = 61.7%, OR = 1.73, 95% CI [1.25, 2.38],  $p = 8.9E^{-4}$ ). A weaker trend of enhanced preferences was found for low-value Go items (mean proportion = 55.6%, OR = 1.30, 95% CI [0.95, 1.77],  $p = .096$ ). As in previous studies using snack-food stimuli<sup>11,12,15</sup>, training with a visual cue resulted in a more robust preference modification for high-value snack-food items than for low-value items (OR = 1.29, 95% CI [1.12, 1.49],  $p = 3.2E^{-4}$ ).

Using an aversive auditory cue in Experiment 6 resulted in a significant (yet not as strong) enhanced preference for high-value Go items (mean proportion = 58.4%, OR = 1.46, 95% CI [1.04, 2.07],  $p = .030$ ) and a significant enhanced preference for low-value Go items (mean proportion = 61.2%, OR = 1.68, 95% CI [1.21, 2.32],  $p = .002$ ), with a more robust

preference change for low-value items compared with high-value items (OR = 0.85, 95% CI [0.73, 0.98],  $p = .022$ ).

### Experiments 7-10: Reproducibility of Findings

After showing CAT can be used to modify preferences for various non-consumable stimuli, to support the reproducibility of our findings, we aimed to replicate our results in four additional independent samples with improved designs.

In Experiment 7 we performed an additional CAT experiment with face stimuli, obtained from a more contemporary dataset<sup>32</sup>, than the stimuli used in Experiment 1. Cue approach training with unfamiliar faces resulted in enhanced preferences for high-value Go items over high-value No-Go items (mean proportion = 69.8%, OR = 2.57, 95% CI [1.86, 3.55],  $p = 8.8E^{-9}$ ); as well as low-value Go items over low-value No-Go items (mean proportion = 68.5%, OR = 2.98, 95% CI [1.67, 5.32],  $p = 2.3E^{-4}$ ). No differences were found between preferences of Go items between high-value and low-value probe choices (OR = 1.08, 95% CI [0.92, 1.25],  $p = .345$ ).

In Experiment 8, CAT training with fractal art images replicated the results of Experiment 2. Participants preferred the high-value Go items over No-Go items (mean proportion = 62.0%, OR = 1.87, 95% CI [1.23, 2.83],  $p = .003$ ); as well as the low-value Go items over the low-value No-Go items (mean proportion = 66.9%, OR = 2.52, 95% CI [1.59, 4.00],  $p = 8.5E^{-9}$ ). The preference change effect was more prominent in low-value choices compared with high-value (OR = 0.78, 95% CI [0.68, 0.91],  $p = 0.001$ ).

In Experiments 9, the association of positive IAPS stimuli with a neutral visual cue resulted in enhanced preferences for Go items, both in the high-value items (mean proportion = 61.6%, log-odds = 1.64, 95% CI [1.35, 2.02],  $p = 1.4E^{-6}$ ) and in the low-value probe choices (mean proportion = 66.7%, log-odds = 2.10, 95% CI [1.65, 2.67],  $p = 1.2E^{-9}$ ). Differences between high and low value probe choices were more robust for low-value items (OR = 0.79, 95% CI [0.70, 0.90],  $p = 5.3E^{-4}$ ).

In Experiments 10, conducted with negative IAPS stimuli and a visual cue, results did not show enhanced preferences for Go items, neither in the high-value pairs (mean proportion = 55.8%, OR = 1.35, 95% CI [0.93, 1.96],  $p = .116$ ), nor in the low-value pairs (mean proportion = 53.3%, OR = 1.20, 95% CI [0.84, 1.71],  $p = .322$ ). No significant differences between high and low value probe choices were found (OR = 1.12, 95% CI [0.98, 1.28],  $p = 0.104$ ).

### **Long Term Maintenance**

In order to evaluate long-term durability of the CAT effect on preferences, all participants from Experiments 2 and 5-8 were encouraged to return for an additional follow-up session, performed one to six months following training (see Table 1 for retention rates). The results of the follow up experiments are listed by the increasing duration of the follow up interval from one to six months.

In a one-month (mean interval 33.2 days) follow-up of Experiment 7 with face stimuli, enhanced preferences for Go items were maintained both in the high-value (mean proportion = 66.8%, OR = 2.23, 95% CI [1.51, 3.27],  $p = 4.8E^{-5}$ ) and the low-value probe choices (mean proportion = 62.1%, OR = 2.16, 95% CI [1.09, 4.27],  $p = .027$ ; see Figure 4 and Table 2). The effect was more robust for high-value than low-value probe choices (OR = 1.26, 95% CI [1.06, 1.50],  $p = 0.007$ ).

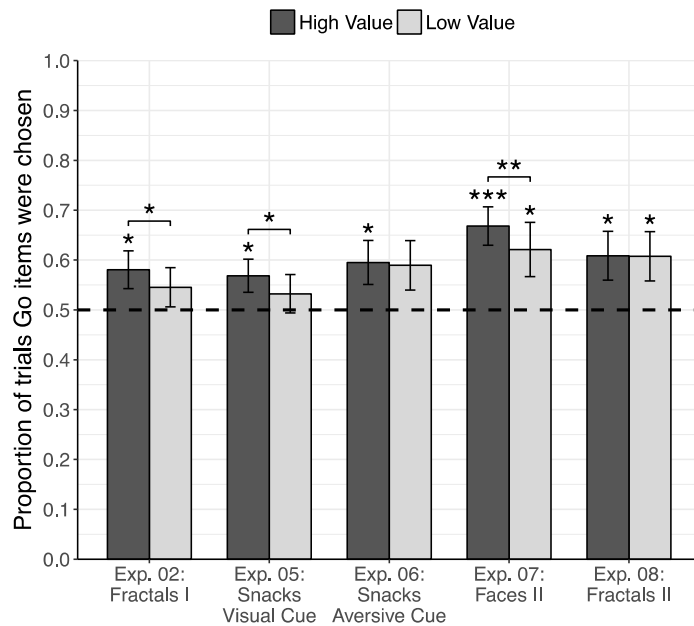
In Experiment 5 with snacks and visual cue, results of the follow-up session conducted approximately two months after the initial training (mean interval 51.8 days), showed that enhanced preferences for the high-value Go items persisted (mean proportion = 56.9%, OR = 1.34, 95% CI [1.02, 1.78],  $p = .038$ ), while preferences for low-value Go items remained at chance level (mean proportion = 53.2%, OR = 1.19, 95% CI [0.85, 1.65],  $p = .317$ ), as in the immediate probe session. The more robust preference for high-value snack-food items found in the first session, were also sustained in the follow-up session (OR = 1.16, 95% CI [1.00, 1.35],  $p = .049$ ).

In Experiment 6 conducted with snack-food stimuli and the aversive auditory cue, following a mean period of two months from training (mean interval 61.5 days), preferences for Go items were maintained in high-value probe choices (mean proportion = 59.4%, OR = 1.52, 95% CI [1.04, 2.24],  $p = .031$ ); while in the low-value probe choices, we observed only a trend of enhanced preferences for Go items (mean proportion = 59.0%, OR = 1.53, 95% CI [0.97, 2.42],  $p = .066$ ), with no difference between high-value and low-value probe choices (OR = 1.02, 95% CI [0.85, 1.23],  $p = .780$ ).

In a four-month (mean interval 116.2 days) follow-up session of Experiment 2 with fractal art stimuli, enhanced preferences were maintained for the high-value Go items (mean proportion = 58.1%, OR = 1.44, 95% CI [1.04, 1.99],  $p = .029$ ), but not for the low-value Go items (mean proportion = 54.5%, OR = 1.23, 95% CI [0.89, 1.71],  $p = .210$ ). Results were more robust for high-value probe choices than low-value probe choices (OR = 1.16, 95% CI [1.02, 1.32],  $p = 0.026$ ).

In a six-month (mean interval 183.9 days) follow-up of Experiment 8 with fractal art stimuli, behavioral change was maintained in both value categories, as participants consistently preferred both the high-value Go items (mean proportion = 60.8%, OR = 1.62, 95% CI [1.05, 2.51],  $p = .030$ ), as well as the low-value Go items (mean proportion = 60.8%, OR = 1.62, 95% CI [1.04, 2.50],  $p = .032$ ). No difference was found between high and low value probe choices (OR = 1.01, 95% CI [0.85, 1.20],  $p = .917$ ).





*Figure 4.* Probe results of follow up sessions.

Results from Experiments 2 (116.2 days after training), Experiment 5 (51.8 days), Experiment 6 (183.9 days), Experiment 7 (33.2 days) and Experiment 8 (183.9 days). Mean proportion of trials participants chose Go items over No-Go items, for high-value (dark gray) and low-value (light gray) probe pairs. Dashed line indicates 50% chance level, error bars represent standard error of the mean. Asterisks reflect statistical significance in a two-tailed logistic regression analysis. Asterisks above each bar represent proportions different from chance (log-odds = 0, odds-ratio = 1). Asterisks above pairs of bars represent differential effect for the two value categories (log-odds = 0, odds-ratio = 1) \*\*\* < 001 \*\* < 01 \* < 05

## Discussion

The current research studies the mechanism and boundary conditions of the novel non-externally reinforced cue-approach task. In 10 independent samples with 255 participants, we showed that cue-approach training can be used to enhance preferences for non-consumable stimuli including faces, fractal art images and positive affective stimuli. We showed that a visual cue can be used to induce the effect as well as an aversive tone with appetitive snack-food items. As a boundary condition we found, in two independent samples, that preferences towards negative affective stimuli could not be changed. In a series of five follow up sessions, performed one to six months following training, we found that the single 40-minutes training session led to a long-lasting preference change. These findings shed light on the mechanism as well as boundaries of this non-reinforced behavior change paradigm.

In this study, we found that CAT was effective in increasing preferences for unfamiliar faces, fractals and positive emotional IAPS images. In the first faces experiment (Experiment 1) we found an inconclusive trend of preference change for stimuli of lower quality and presumably a too-wide value range; therefore, in an improved replication (Experiment 7) we changed the stimuli and narrowed the value category. This induced a stronger and more consistent preference change. The ability to change preferences for abstract stimuli, such as fractals (Experiments 2 and 8), exemplifies the generalizability of the non-reinforced behavioral change mechanism underlying CAT, which goes beyond stimuli of consumable, familiar and even tangible nature.

Using affective stimuli in Experiments 3-4 and 9-10, we found that cue approach can modify preferences for stimuli of positive valence (Experiments 3 and 9). However, we were not able to induce this behavioral change for negative affective stimuli, neither with an auditory cue nor with a visual cue (Experiments 4 and 10, respectively). Replicating these null results with a visual cue assisted in ensuring that participants were looking at the stimuli during training. This was especially important in experiments with negative affective stimuli, where participants might have been motivated to avert their gaze from the unpleasant stimuli during training, which could provide alternative explanation for the null results. These findings suggest an important boundary condition of the CAT paradigm, showing that the association of a Go cue can be used to change preferences for positive but not negative affective stimuli. These findings correspond with similar results with operant conditioning, which showed that go responses were better learned by association with rewards, while no-go responses were better learned by association with punishments<sup>19-21</sup>. Future studies could test this hypothesis using a non-reinforced paradigm that would require a no-go response, in order to modify preferences for negative affective stimuli.

Using familiar local snack-food items with a neutral visual cue and an auditory aversive cue (Experiments 5-6), we found that the CAT effect is not limited to a specific cue

modality. In Experiment 5, we showed that using a visual cue rather than the original neutral tone led to similar results as in previous CAT studies with local snacks<sup>11,12</sup>. These results suggest that modifying the modality of the cue from auditory to visual does not impair the behavioral change effect. Interestingly, an association with an aversive auditory cue (Experiment 6), which could have been expected to decrease preferences for associated items via classical conditioning<sup>25</sup>, resulted in enhanced preferences for both high and low value Go items. Using an aversive cue resulted in a significantly stronger preference modification for the less liked low-value snack-food items. This was surprising given that in other CAT experiments conducted with snack-food items there was no change or a weaker change for low-value food items<sup>11,12</sup>. This might suggest an important interaction between stimuli and cue, as a less-positive Go cue may be more effective in modifying preferences for less-favorable stimuli. Future research could directly test the effect of high versus low valued stimuli using additional aversive cues, such as unpleasant tactile or electric stimulation.

In previous studies conducted with snack-food items<sup>11,12</sup>, preference change was shown to have a differential effect as a function of value category, with robust changes of preferences for high-value snacks, yet weaker changes for low-value snacks. However, in the current investigation, in samples using non-consumable items (faces, fractals and affective stimuli) we did not observe such a consistent differential effect in each experiment. This might suggest that the greater change for higher-valued items is linked to previous experience, such as existing preferences for familiar snack-food items. Our results imply that the mechanisms underlying CAT may be used to modify preferences for the entire neutral and positive value range of unfamiliar stimuli, for which participants had no previous experience. Our null results with the negative affective stimuli experiments, may resemble the differential effect for high-value versus low-value snack-food items, as in both cases training the less positive stimulus did not induce an enhancement of preferences. These findings hedge the boundaries of the learning mechanism underlying the CAT, suggesting it

may be specific to neutral and positive rather than negative stimuli, even when using an aversive cue. The consistent finding that the task could not modify preferences for less liked snack-food items in previous studies, as with the two experiments using negative affective stimuli, also serves as evidence against a concern that the experimental results may stem from participants trying to affirm the researchers' aim due to demand characteristics<sup>43</sup>. Taken together, the lack of ability to change preferences for lower value items suggests a potential thresholding mechanism, such that CAT can only affect items above a certain initial preference level<sup>44</sup>.

In five of the experiments, we included a follow-up session, one to six months after training. In all cases, the behavioral change effect persisted over time. These results point to the high durability of behavioral change induced without external reinforcement or context changes. A relatively short 40 minutes training session successfully affected preferences, which then persisted over long periods of time, up to six months after training, with no maintenance procedures between these time points. This putatively suggests that the training induces a change in low-level brain areas<sup>29</sup> that lingers for several months. Thus, this non-reinforced mechanism could be valuable for designing effective real-world interventions that could induce a long-lasting impact.

The work presented here serves as a foundation for studies examining the real-world applications of the CAT. Applicable implementation of the CAT can be used to enhance desired behaviors over less desired ones, not necessarily limited to food consumption. In the clinical field, several psychological disorders such as depression and anxiety are characterized by a cognitive bias towards negative affective social stimuli<sup>45,46</sup>. Attentional bias modification treatments for these disorders have been tested in an attempt to improve clinical symptoms<sup>47,48</sup>. Similarly, the general mechanism underlying CAT may be used to induce a counter-bias and enhance preferences for positive affective stimuli. Such preference

modification may, in turn, lead to congruent changes in positive mood with beneficial long term therapeutic effects.

In conclusion, in the current work, we show that the cue approach task, previously shown to change preferences for consumable food items, has a wide reach beyond specific stimuli and cues. We demonstrate that the behavioral change induced by a short 40-minute training procedure is durable over long periods of several months after training. Our findings suggest that non-externally reinforced behavioral change holds great potential to develop novel applications with long term efficacy, which can be used to enhance desired behavior in a wide array of domains.

## References

1. Butler, A. C., Chapman, J. E., Forman, E. M. & Beck, A. T. The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clin. Psychol. Rev.* **26**, 17–31 (2006).
2. Higgins, S. T. *et al.* Outpatient behavioral treatment for cocaine dependence: One-year outcome. *Exp. Clin. Psychopharmacol.* **3**, 205–212 (1995).
3. Lang, W., Wing, R. R., Venditti, E., Jakicic, J. M. & Polley, B. A. Lifestyle Intervention in Overweight Individuals With a Family History of Diabetes. *Diabetes Care* **21**, 350–259 (1998).
4. Bjork, R. A. Recency and Recovery in Human Memory. *Nat. Rememb. Essays Honor Robert G. Crowder. Sci. Conf. Ser.* 211–232 (2001). doi:10.1037/10394-011
5. Bouton, M. E. Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychol. Bull.* **114**, 80–99 (1993).
6. Cahill, K. & Perera, R. Competitions and incentives for smoking cessation. Cochrane Database of Systematic Reviews. Issue 4. Art. No.: CD004307 (2011). doi:10.1002/14651858.CD004307.pub4.Copyright
7. Wood, W. & Neal, D. T. A new look at habits and the habit-goal interface. *Psychol. Rev.* **114**, 843–863 (2007).
8. De Martino, B., Kumaran, D., Seymour, B. & Dolan, R. J. Frames, Biases, and Rational Decision Making in the Human Brain. *Science (80-. )*. **313**, 684–687 (2006).
9. Slovic, P. The construction of preference. *Am. Psychol.* **50**, 364–371 (1995).
10. Tversky, A. & Kahneman, D. Rational Choice and the Framing of Decisions. *J. Bus.* **59**, 251–278 (1986).
11. Schonberg, T. *et al.* Changing value through cued approach: an automatic mechanism of behavior change. *Nat. Neurosci.* **17**, 625–30 (2014).
12. Bakkour, A. *et al.* Mechanisms of choice behavior shift using cue-approach training. *Front. Psychol.* **7**, 1–12 (2016).
13. Veling, H. *et al.* Training Impulsive Choices for Healthy and Sustainable Food. *J. Exp. Psychol. Appl.* **23**, 204–215 (2017).
14. Zoltak, M. J., Veling, H., Chen, Z. & Holland, R. W. Attention ! Can choices for low value food over high value food be trained? *Appetite* 1–9 (2017). doi:10.1016/j.appet.2017.06.010
15. Bakkour, A., Lewis-peacock, J. A., Poldrack, R. A. & Schonberg, T. Neural mechanisms of cue-approach training. *Neuroimage* **151**, 92–104 (2017).

16. Swallow, K. M. & Jiang, Y. V. The Attentional Boost Effect: Transient increases in attention to one task enhance performance in a second task. *Cognition* **115**, 118–132 (2010).
17. Aharon, I. *et al.* Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* **32**, 537–551 (2001).
18. Kranz, F. & Ishai, A. Face perception is modulated by sexual preference. *Curr. Biol.* **16**, 63–68 (2006).
19. Guitart-Masip, M. *et al.* Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J. Neurosci.* **31**, 7867–7875 (2011).
20. Guitart-Masip, M. *et al.* NeuroImage Go and no-go learning in reward and punishment : Interactions between affect and effect. *Neuroimage* **62**, 154–166 (2012).
21. Chiu, Y., Cools, R. & Aron, A. R. Opposing Effects of Appetitive and Aversive Cues on Go / No-go Behavior and Motor Excitability. *J. Cogn. Neurosci.* **26**, 1851–1860 (2014).
22. Lang, P. J., Bradley, M. M. & Cuthbert, B. N. *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8* (2008).
23. Ng, A. W. Y. & Chan, A. H. S. Finger Response Times to Visual, Auditory and Tactile Modality Stimuli. *Proc. Int. MultiConference Eng. Comput. Sci.* **II**, 1449–1454 (2012).
24. Stein, B., London, N., Wilkinson, L. & Price, D. Enhancement of Perceived Visual Intensity by Auditory Stimuli: A Psychophysical Analysis. *J. Cogn. Neurosci.* **8**, 497–506 (1996).
25. Neumann, D. L. & Waters, A. M. The use of an unpleasant sound as an unconditional stimulus in a human aversive Pavlovian conditioning procedure. *Biol. Psychol.* **73**, 175–185 (2006).
26. Ioannidis, J. P. A. Why most published research findings are false. *PLoS Med.* **2**, 696–701 (2005).
27. Munafò, M. R. *et al.* A manifesto for reproducible science. *Nat. Hum. Behav.* **1**, (2017).
28. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science (80-. ).* **349**, (2015).
29. Serences, J. T. & Yantis, S. Selective visual attention and perceptual coherence. *Trends Cogn. Sci.* **10**, 38–45 (2006).

30. Champely, S. pwr: Basic Functions for Power Analysis. (2017).
31. Yovel, G. & Kanwisher, N. Face Perception : Domain Specific, Not Process Specific. *Neuron* **44**, 889–898 (2004).
32. Vieira, T. F., Bottino, A., Laurentini, A. & De Simone, M. Detecting siblings in image pairs. *Vis. Comput.* **30**, 1333–1345 (2014).
33. Fantastic Fractals. (2013). Available at: <http://fantastic-fractals.com>.
34. Snack-Food Items Dataset. (2017). Available at: <http://schonberglab.tau.ac.il/resources/snack-food-image-database/>.
35. Bigman, Y. E., Sheppes, G. & Tamir, M. When Less is More: Effects of the Availability of Strategic Options on Regulating Negative Emotions. *Emotion* (2017).
36. Kleiner, M. *et al.* What's new in Psychtoolbox-3. *Perception* **36**, 1 (2007).
37. Shinnars, P. PyGame–python game development. (2011). Available at: <http://www.pygame.org/>.
38. Becker, G. M., DeGroot, M. H. & Marschak, J. Measuring utility by a single-response sequential method. *Behav. Sci.* **9**, 226–232 (1964).
39. Von Neumann, J. & Morgenstern, O. *Theory of games and economic behavior*. (Princeton university press, 1944).
40. Colley, W. Colley's bias free college football ranking method: The colley matrix explained. *Princet. Univ.* 1–23 (2002).
41. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4 Douglas. *J. Stat. Softw.* **67**, 1–48 (2015).
42. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag, 2009).
43. Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *Am. Psychol.* **17**, 776–783 (1962).
44. McGuire, J. T. & Kable, J. W. Go means green. *Nat. Neurosci.* **17**, 489–490 (2014).
45. Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J. & van IJzendoorn, M. H. Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol. Bull.* **133**, 1–24 (2007).
46. Surguladze, S. *et al.* A differential pattern of neural response toward sad versus happy facial expressions in major depressive disorder. *Biol. Psychiatry* **57**, 201–209 (2005).
47. Browning, M., Holmes, E. A., Charles, M., Cowen, P. J. & Harmer, C. J. Using attentional bias modification as a cognitive vaccine against depression. *Biol. Psychiatry* **72**, 572–579 (2012).



48. Hakamata, Y. *et al.* Attention bias modification treatment: A meta-analysis toward the establishment of novel treatment for anxiety. *Biol. Psychiatry* **68**, 982–990 (2010).