**Preregistration questions for OSF**

**Description (optional)**

**Please give a brief description of your study, including some background, the purpose of the study, or broad research questions.**

The Cue Approach Training (CAT) has been recently introduced as a novel paradigm to modify preferences without external reinforcements, via mere association of image stimuli with a cue and rapid response (Schonberg et al., 2014). The experimental procedure of CAT includes three main phases: initial preferences evaluation task, CAT task, and probe task. In the initial preference evaluation task, using an auction procedure (Becker, DeGroot, & Marschak, 1964) or forced binary choice, the subjective preferences of each participant are evaluated for a set of stimuli. In the CAT task, stimuli are presented one at a time. Approximately 30% of stimuli are presented with a delayed cue to which participants need to respond consistently with a rapid button press response (Go stimuli). The rest of the stimuli appear without being associated with a cue and response (NoGo stimuli). In the probe phase, preference modification effect is evaluated in a forced choice task. In each trial, two stimuli of similar initial subjective value are pitted against each other (both either high-value or low-value). In each pair, one stimulus is a Go stimulus, and the other NoGo stimulus.

In dozens of experiments, using different stimuli (snacks, faces, fractals, positive affective stimuli) and cues (auditory, visual, aversive auditory) CAT was found to modify preferences, as participants consistently preferred the associated Go items over the NoGo items (Bakkour et al., 2016; Bakkour, Lewis-Peacock, Poldrack, & Schonberg, 2017; Botvinik-Nezer, Salomon, & Schonberg, 2019; Salomon et al., 2018, 2019; Schonberg et al., 2014; Zoltak, Veling, Chen, & Holland, 2018).

While preference modification following CAT was well founded as a stable group-level effect, a great deal of variability was observed across participants, with some participants showing strong preference modification effect and others weaker one. In the current work we aimed to identify a computational marker for individual differences in CAT learning.

In the current study we will examine reaction time (RT) patterns, and using a Bayesian modelling approach, we aim to identify a computational marker for learning. We found unique patterns associated with the training repetitions in the task – while in early training runs, RTs were characterized with a narrow unimodal distribution following the Cue onset (Cue-dependent responses), as training progressed, the RT distribution shifted towards an earlier left-tail distribution of early responses, sometimes preceding the actual cue onset (anticipatory responses). To model the RT dynamics, we use a Bayesian computational model approach. RT distributions were modelled as mixture of two Gaussians with different means and standard deviations. The mixture proportion () was defined as a time-dependent parameter. A participant-level random slope () determined the transition rate from late cue-dependent RT to early anticipatory RT, as a function of the run.

The random slope () parameter will serve in the current work as potential computational marker for learning, which could account for individual differences in task performance between participants. In the current work we will examine the association between the marker and preference modification effect following CAT. We will also aim to manipulate the via a novel CAT design. We intend to manipulate the parameter estimate within-participant by associating half of the Go stimuli with the cue at 100% contingency (similar to previous CAT designs), and half of the Go stimuli with only 50% contingency (i.e. cue will be associated with the stimuli in only 50% of training trials).

**Hypotheses (required)**

**List specific, concise, and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here. If a specific interaction or moderation is important to your research, you can list that as a separate hypothesis.**

**Probe analysis**: In line with previous findings with CAT, we expect participants will choose Go stimuli over similar value NoGo stimuli above chance level (50% proportion; log-odds = 0; odds-ratio = 1). Thus, the CAT preference modification effect will be tested using a one-sided repeated measures logistic regression analysis.

We will test the effect across all value categories combined (i.e. higher Go versus NoGo choices will be analyzed together with lower Go versus NoGo choices). We do not expect to find a significantly stronger effect for probe pairs of different value categories (Salomon et al., 2018); therefore, the effect within value categories will be evaluated using an hierarchical two-sided logistic regression model.

We also expect that preference modification will be more robust for the 100% contingency Go stimuli, compared to 50% contingency ones. Therefore, we will test preference modification in a one-sided hierarchical logistic regression model – first examining the preference modification separately for each contingency, followed by an additional analysis with the contingency factor examining whether the 100% contingency resulted in greater preference modification to Go stimuli, compared with the 50% contingency.

**Bayesian model for CAT**: The main goal of the current work, will be to establish a computational marker for non-reinforced learning. For that aim, we utilize a Bayesian modelling approach. Our second analysis of interest will examine the random parameter, contrasting the two contingency conditions - We hypothesize that 100% contingency Go association would result in more robust parameter estimates than 50% contingency. We will contrast these values within participants using mixed linear regression (with participants as random effect).

Finally, our third analysis would examine the parameter as a computational marker for value. We will examine whether individual differences in the parameter are indicative of individual differences in preference modification effect. For each participant we will calculate the proportion of trials in which Go items were chosen during the probe phase as our outcome variable and use parameter as a predictor for this behavioral change effect, using one sided linear mixed model. We will perform this analysis separately for 100% contingency Go stimuli and 50% contingency.

**Study design (required)**

**Describe your study design. Examples include two-group, factorial, randomized block, and repeated measures. Is it a between (unpaired), within-subject (paired), or mixed design? Describe any counterbalancing required. Typical study designs for observation studies include cohort, cross sectional, and case-control studies.**

**Materials**

**Stimuli.** In the current project, we will use a stimulus set of 80 face images, adapted from the Siblings Dataset (Vieira, Bottino, Laurentini, & De Simone, 2014), as used in a previous behavioral CAT publication with faces (Salomon et al., 2018, 2019). The stimulus set comprises of 40 male and 40 female front-facing individuals, posing a neutral expression with limited facial hair and make-up. The original images were cropped to identical size (400 × 500 pixels) and the original green screen background was replaced with a homogenous gray background. Faces were aligned by positioning each figure’s pupils in fixed coordinates symmetrically around the center of the image ([150, 250] and [250, 250] for the left and right figure’s pupil, respectively).

**Cue.** In the training task, we will use a neutral visual cue of a 96x96 pixels semitransparent Gabor image, presented for 100ms on top of the face stimuli.

Procedure

Overall procedure will follow a similar course to that of previous CAT studies with faces (Salomon et al., 2018, 2019) – an initial preference evaluation task, followed by CAT and a binary choice probe phase, with most prominent modifications in the CAT task.

Baseline evaluation of subjective preference. Participants’ baseline subjective preference for the 80 individual face stimuli will be evaluated using a forced-choice binary ranking procedure. Participants will be presented with 400 unique binary choices in which they will be asked to select their preferred stimulus out of two randomly paired face stimuli, within an allocated 2000-ms time window. Each stimulus will be presented in exactly 10 choice trials to maintain a similar exposure to all stimuli

Similar to our previous work with face stimuli (Salomon et al., 2018), binary choices acquired in the baseline subjective preference phase, will be transformed into ranking scores using the Colley Matrix algorithm (Colley, 2002). Stimuli will be rank-sorted according to their initial value and the following CAT and probe task will be based on these rankings. The 80 stimuli will be than categorized into 10 value groups of similar value, each with 8 stimuli (i.e. – highest value will be of stimuli ranked 1-8, followed by stimuli ranked 9-16, and so forth until the lowest value-category of stimuli ranked 83-90).

**Cue-approach training.** The CAT task’s protocol will follow that of previous CAT with faces (Salomon et al., 2018), with modifications aimed to optimize the experiment for Bayesian modelling of RT and manipulating the learning parameter. Like previous experiments, each stimulus in the training set will be presented individually on the screen for 1000-ms in each trial, once during each of the 20 training runs. During Go trials a visual cue in the form of semi-transparent Gabor will appear on top of the image stimuli. Participants will be required to respond with a rapid button in each Go trial. During NoGo trials, a visual cue will not be presented and participants will not be required to perform a response. Participants will not receive any feedback about their performance throughout the task, neither after individual trials, nor at the end of the task.

Unlike previous CAT experiments, which employed adaptive ladder procedure to adjust difficulty to participants’ performance, in the current experiment, in Go trials the visual cue will always be presented after a fixed interval of 850ms following the trial onset. This will optimize the efficiency of the model, as well as encourage participants to generate anticipatory responses in order to press on time, before trial offset.

An additional modification would be the introduction of two classes of Go stimuli. Half of the Go stimuli will be associated with the Go signal in 100% of the trials they appear in (as in previous CAT studies; 100% contingency), while the other half of Go stimuli will be associated with the cue only in 50% of the time (10 of the 20 training runs; 50% contingency). Overall Go trials proportion is balanced across training runs. Unlike previous CAT experiments, participants will be clearly disclosed of this association as part of the task’s instructions - participants will be explained that some stimuli are always presented with a cue, some only in 50% of trials will be presented with a cue, and some will never be presented with a cue. Participants will also be encouraged to produce anticipatory responses by instructing them to respond when the cue appears or when they know that the cue will appear, while maintaining their responses as fast and as accurate as they can (i.e. avoid false alarm).

Go allocation is decided based on initial subjective preferences, so that Go and NoGo stimuli, as well as 50% Go to 100% Go associations, are all matched on initial-value. The highest and lowest value categories (ranks 1-8 and 73-80) are all allocated to be NoGo stimuli. In each value category, 4 out of the 8 stimuli are Go and the rest NoGo, so that the mean rank of Go and NoGo is matched - e.g. for the value category with the stimuli ranked 9-16, for Go allocation stimuli 9, 12, 13, and 16 would be selected (mean rank = 12.5 both for Go and NoGo stimuli). Similar method will be used to balance the allocation to 100% versus 50% association stimuli. In total, 32 NoGo stimuli (four stimuli in eight of the value-categories) are value-matched with 32 Go stimuli. Of the Go stimuli, 16 are associated with the cue with 100% contingency, and 16 with 50% contingency. Throughout the entire task, 30% of all trials require Go response, similarly to previous CAT experiments.

**Probe.** Similar to previous CAT experiments, in the probe phase participants will perform binary choices, choosing their preferred stimulus from pairs of stimuli. In each pair, a Go stimulus (either 50% Go contingency or a 100% contingency) will be pitted against a NoGo stimulus of similar initial value (same value category). Preferences modification will be evaluated as deviation from the expected 50% chance of choosing Go stimuli over NoGo stimuli. For each of the eight value categories, the four Go stimuli will be pitted against the four NoGo stimuli, resulting in 16 (4x4) probe pairs per value-category, and 128 total probe trials – 64 trials of 100% Go versus NoGo and 64 trials of 50% Go versus NoGo.

**Randomization (optional)**

**If you are doing a randomized study, how will you randomize, and at what level?**

In all tasks stimuli presentation order will be randomized. Go allocation (i.e. which stimuli will be associated with the Go cue at 100% contingency, 50% contingency or none at all) will be randomized while maintaining equal mean value. Overall Go trials proportion is balanced across training runs.

**Data collection procedures (required)**

**Please describe the process by which you will collect your data. If you are using human subjects, this should include the population from which you obtain subjects, recruitment efforts, payment for participation, how subjects will be selected for eligibility from the initial pool (e.g. inclusion and exclusion rules), and your study timeline. For studies that don't include human subjects, include information about how you will collect samples, duration of data gathering efforts, source or location of samples, or batch numbers you will use.**

The study will include *n* = 59 valid participants, ages 18-39, with correct or corrected to normal vision. The study sample size was determined based on a preliminary pilot study, on which we performed 95% power analysis (see further details in ‘Power analysis’ section below). Participants will give their informed consent to participate in the experiment and received monetary compensation for their time. The study is approved by the ethics committee of Tel Aviv University.

**Exclusion criteria.** Participants will be excluded from the analysis based on the following three criteria, adapted based on previous CAT studies (Botvinik-Nezer et al., 2019; Salomon et al., 2018, 2019; Schonberg et al., 2014):

*Initial preference intransitivity*. Intransitive choice pattern results in dense distribution of ranking values using the Colley ranking algorithm (Colley, 2002), characterized with small variability in ranking-scores of the 80 stimuli. The standard deviation of Colley scores will be used as a transitivity score. Outliers with very small transitivity score (Z < -3, below the group mean) will be excluded.

*False alarm*. During CAT we will evaluate the mean proportion of trials in which a false response was made to a NoGo stimulus. Participants with high rate of false alarm (Z > 3, above the group mean) will be excluded.

*Miss*. During CAT we will evaluate the standard proportion of trials no response was made to Go stimuli. Participants with high rate of missed trials (Z > 3, above the group mean) will be excluded

**Sample size (required)**

**Describe the sample size of your study. How many units will be analyzed in the study? This could be the number of people, birds, classrooms, plots, interactions, or countries included. If the units are not individuals, then describe the size requirements for each unit. If you are using a clustered or multilevel design, how many units are you collecting at each level of the analysis?**

The study will include *n* = 59 valid participants, ages 18-39, with correct or corrected to normal vision. The study sample size was determined based on a preliminary pilot study, on which we performed 95% power analysis.

**Sample size rationale (optional)**

**This could include a power analysis or an arbitrary constraint such as time, money, or personnel.**

To determine the sample size for the next pre-registered experiment, we performed a power analysis based on the pilot experiment results. We tested four effects of interest – an effect of enhanced preference for Go over NoGo stimuli, using logistic regression (two effects, for the 100% and 50% contingency conditions) and a correlation between each of the two θslope parameter estimates and CAT preference modification effect, using linear regression. Since θslope parameter estimates and logistic regression effect cannot be analytically solved, we used a bootstrapping resampling approach to estimate the power for different sample size. For each sample size, we resampled (with replacement) 1000 random samples with from our pilot study data. We estimated power as the proportion of samples (out of 1000) in which the analysis acceded statistical significance threshold (𝛼 = .05, two-sided). To account for near-threshold fluctuations, power was tested for five subsequent sample sizes before terminating the power analysis. All data and R codes are shared in this OSF project.

To account for underestimation of variability due to resampling of the same participants, we used a stringent power threshold of 95% to determine the sample size for the full experiment. Our analysis suggests that to achieve 95% power to detect all four effects of interest, we would need to collect a sample of *n* = 59 participants.

**Manipulated variables (optional)**

**Describe all variables you plan to manipulate and the levels or treatment arms of each variable. This is not applicable to any observational study.**

To manipulate the computational learning parameter within participant we will use two classes of Go stimuli. Half of the Go stimuli will be associated with the Go signal in 100% of trials (as in previous CAT studies; 100% contingency), while the other half of Go stimuli will be associated with the cue 50% of the time (10 of the 20 training runs; 50% contingency). Overall Go trials proportion is balanced across training runs. Unlike previous CAT experiments, participants will be clearly disclosed of this association as part of the task’s instruction - they will be explained that some stimuli are always presented with a cue, some only in 50% of trials will be presented with a cue, and some will never be presented with a cue. Participants will also be encouraged to produce anticipatory responses by instructing them to respond when the cue appears or when they know that the cue will appear, while maintaining their responses as fast and as accurate as they can (i.e. avoid false alarm).

**Measured variables (required)**

**Describe each variable that you will measure. This will include outcome measures, as well as any predictors or covariates that you will measure. You do not need to include any variables that you plan on collecting if they are not going to be included in the confirmatory analyses of this study.**

In the probe phase, the variable of interest will be a binary outcome – whether in each trial participants chose the Go stimulus or the NoGo. This outcome variable will be both analyzed as a binary outcome in logistic regression analyses, as well as transformed into mean proportion of trials Go items were chosen, which will then be regressed with the computational learning parameter.

The main goal of the current work, will be to establish a computational marker for non-reinforced learning. For that aim, we utilize a Bayesian modelling approach. Reaction-time (RT) patterns will be modelled using a Bayesian model (implemented with Stan and RStan package for R). Reaction time will be fitted to a mixture of Gaussians model: one Gaussian distribution of short RT, with a mean preceding the cue onset (fixed at 850 ms), and a second distribution of mean above cue onset. The mixture proportion , will be modelled (as detailed in the full document attached to this preregistration), using a fixed and a parameter, evaluated for each participant. The parameter will account for the rate of transition from cue-dependent responses to early anticipatory responses (learning parameter of interest). The will also be evaluated separately for 100% contingency Go stimuli () and 50% contingency Go stimuli ().

The parameter will be our outcome variable of interest from the CAT task. We will compare the estimated parameter between contingency levels, as well as examine its correlation with the probe behavioral effect.

**Statistical models (required)**

**What statistical model will you use to test each hypothesis? Please include the type of model (e.g. ANOVA, multiple regression, SEM, etc) and the specification of the model (this includes each variable that will be included as predictors, outcomes, or covariates). Please specify any interactions, subgroup analyses, pairwise or complex contrasts, or follow-up tests from omnibus tests. If you plan on using any positive controls, negative controls, or manipulation checks you may mention that here. Remember that any test not included here must be noted as an exploratory test in your final article.**

**Probe analysis**. Similar to previous experiments with CAT, we will use logistic regression to evaluate preference modification following CAT. In line with previous work using CAT and faces (Salomon et al., 2018, 2019), we expect participants to choose Go stimuli over similar value NoGo stimuli above chance level (50% proportion; log-odds = 0; odds-ratio = 1). Thus, CAT preference modification effect will be tested using a one-sided repeated measures logistic regression analysis. We will test the effect across all value categories combined. We do not expect to find a significantly stronger effect for probe pairs of different value categories (Salomon et al., 2018); therefore, value category will be evaluated using an hierarchical two-sided logistic regression model.

We also expect that preference modification will be more robust for the 100% contingency Go stimuli, compared to 50% contingency ones. Therefore, we will test preference modification using one-sided hierarchical logistic regression model – first examining the preference modification separately for each contingency, followed by an additional analysis with the contingency factor examining whether preference modification was more robust to the 100% contingency stimuli.

**Bayesian models for CAT**. The main goal of the current work, will be to establish a computational marker for non-reinforced learning. For that aim, we utilize a Bayesian modelling approach. Reaction-time patterns for Go trial in the CAT task will be modelled in a Bayesian model (implemented with Stan and RStan package for R). Reaction time distribution will be fitted to a mixture of Gaussians model: one Gaussian distribution of short RT, with a mean preceding the cue onset (fixed at 850 ms), and a second distribution of mean above cue onset. The mixture proportion , will be modelled, as described in the associated document, using a fixed and a parameter, evaluated for each participant. The parameter will account for the rate of transition from late cue-dependent responses to early anticipatory responses (learning parameter of interest). The will also be evaluated separately for 100% contingency Go stimuli () and 50% contingency Go stimuli (). To account for variability of early onset RT between participants, the mean of anticipatory responses (), will also be

Our second analysis of interest will examine the random parameter, contrasting for the two conditions. We hypothesize that 100% contingency Go association would result in more robust parameter estimates than 50% contingency. We will contrast these values within participants using mixed linear regression (with participants as random effect).

Finally, our third and most crucial analysis would use the parameter as computational marker for value. We will examine whether individual differences in parameter are indicative of individual differences in preference modification effect. For each participant we will calculate the proportion of trials Go stimuli were chosen during the probe phase as our outcome variable, and use parameter as predictor for this behavioral change effect, using one sided linear mixed model. We will perform this analysis separately for 100% contingency Go stimuli and 50% contingency.

**Transformations (optional)**

**If you plan on transforming, centering, recoding the data, or will require a coding scheme for categorical variables, please describe that process.**

In the probe phase, the variable of interest will be a binary outcome – indicating in each pair whether participants chose the Go stimulus or the NoGo. This outcome variable will be both analyzed as a binary outcome in logistic regression analyses, as well as transformed into mean proportion of trials Go items were chosen, which will then be regressed with the computational learning parameter.

**Inference criteria (optional)**

**What criteria will you use to make inferences? Please describe the information you’ll use (e.g. specify the p-values, Bayes factors, specific model fit indices), as well as cut-off criterion, where appropriate. Will you be using one or two tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this?**

We will use p-values of 0.05 to determine statistical significance. For directional hypotheses we will report one-sided statistical significance. We will also report measurement for effect size (e.g. OR in logistic regression, regression coefficients for regression models) along with dispersion measurements.

**Data exclusion (optional)**

**How will you determine which data points or samples if any to exclude from your analyses? How will outliers be handled? Will you use any awareness check?**

Exclusion will be done for entire participants’ data, according to the exclusion criteria mentioned above.

**Missing data (optional)**

**How will you deal with incomplete or missing data?**

In case of missing or incomplete data (e.g. participants who fail to make a choice within the allocated time frame), the data will remain missing and will be not be included in the analysis. All statistical tests we use (e.g. logistic regression) can run while excluding the missing data.

**Exploratory analysis (optional)**

**If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time.**

We might decide to modify the details of our Bayesian model to include more sophisticated data (e.g. random slopes for individual stimuli). We will report in the manuscript all deviations from this preregistration, and clarify which analyses were exploratory.