Journal of Experimental Psychology: General

ISSN: 0096-3445

2025, Vol. 154, No. 7, 1888–1915
https://doi.org/10.1037/xge0001739

A Computational Model for Individual Differences in Nonreinforced Learning

Tom Salomon[1], Alon Itzkovitch[1], Nathaniel D. Daw[2], and Tom Schonberg[1, 3]
[1] School of Neurobiology, Biochemistry and Biophysics, Faculty of Life Sciences, Tel Aviv University
[2] Department of Psychology, Neuroscience Institute, Princeton University
[3] Sagol School of Neuroscience, Tel Aviv University

Cue-Approach Training (CAT) is a paradigm that enhances preferences without external reinforcements, suggesting a potential role for internal learning processes. Here, we developed a novel Bayesian computational model to quantify anticipatory response patterns during the training phase of CAT. This phase includes individual items, and thus, this marker potentially reflects internal learning signals at the item level. Our model, fitted to meta-analysis data from 28 prior CAT experiments, was able to predict individual differences in nonreinforced preference changes using a key computational marker. Crucially, two new experiments manipulated the training procedure to influence the model's predicted learning marker. As predicted and preregistered, the manipulation successfully induced differential preference changes, supporting a causal role of our model. These findings demonstrate powerful potential of our computational framework for investigating intrinsic learning processes. This framework could be used to predict preference changes and opens new avenues for understanding intrinsic motivation and decision making.

Public Significance Statement
This study provides new insights into how people can learn and change their preferences without relying on external reinforcements, a process relevant to everyday decision making. By developing a computational model on data from 28 studies with 840 participants, we demonstrate that individual differences in how quickly people learn during a simple task can predict changes in their future preferences. These findings help explain how preferences, such as choosing healthier foods, forming habits, or responding to advertising, can shift without immediate external reinforcements. This model has potential applications in public health, education, and marketing, where understanding and influencing decision making without direct reinforcement can improve outcomes in areas such as healthy living, behavior change, and personalized interventions.

Keywords: cognitive neuroscience, decision making, computational modeling, learning

Supplemental materials: https://doi.org/10.1037/xge0001739.supp

This article was published Online First May 22, 2025.
Rachel Wu served as action editor.
Tom Salomon https://orcid.org/0000-0002-1417-8163
Tom Schonberg https://orcid.org/0000-0002-4485-816X
Tom Salomon and Alon Itzkovitch contributed equally to this work. All materials and analysis codes are available online at https://github.com/tomsalomon/CAT_Individualized_Learning. Preregistration hypotheses, experimental design data, and codes used for power analysis are available in the Open Science Foundation depository (https://osf.io/nwr4v). This article is available as a preprint on the bioRxiv preprint server at https://doi.org/10.1101/2022.03.20.484477.

The authors have no competing interests. Results incorporated here have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 and Horizon Europe research and innovation programs (Grants 715016 and 101087869). The views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the ERC Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. This work was also supported by the Israeli Science Foundation (1798/15 and 1996/20) awarded to Tom Schonberg. Tom Salomon was supported by the Nehemia Levtzion fellowship and the Fields-Rayant Minducate Learning Innovation Research Center. This research was funded in whole, or in part, by the Horizon Europe Framework Programme (Grant 715016 and Grant 101087869). For the purpose of open access, the author has applied a CC BY public copyright license to any Author accepted manuscript version arising from this submission.

Tom Salomon played a lead role in conceptualization, data curation, formal analysis, methodology, visualization, writing–original draft, and writing–review and editing. Alon Itzkovitch played a supporting role in

continued

A core concept is that choice is ultimately driven by preferences over outcomes (Samuelson, 1938). Accordingly, how people construct and learn or modify their evaluations is commonly understood as reflecting processes that ultimately concern learning about those outcomes, that is, learning driven by external reinforcements (Niv, 2009; Rangel et al., 2008; Staddon & Cerutti, 2003; Sutton & Barto, 2018). At the same time, many findings suggest that seemingly reinforcement-irrelevant manipulations can impact choice (Krajbich et al., 2010; Krajbich & Rangel, 2011; Schonberg et al., 2014). Because they do not involve reinforcement, these puzzling findings may help to shed light on the covert processes and representations driving evaluation, which remain largely mysterious. One relevant area of inquiry concerns value modification. While reinforcement-based learning dominated this field, a recent review (Schonberg & Katz, 2020) highlighted various means to influence preferences and choices without external reinforcements covered in the literature, dating back to the *mere exposure effect*. In these studies, simple repeated exposure to stimuli was found to enhance preferences (Zajonc, 1968). Similarly, studies that demonstrated preferences can be shaped by previous choices (Voigt et al., 2017), provide another example of nonreinforced preference change paradigm. These studies provide a strong theoretical framework for the efficacy of interventions such as advertisements on preferences and how life choices affect future decisions (Sharot et al., 2010, 2012). Because they go beyond external reinforcement-based manipulations (Rindfleisch & Inman, 1998; Tom et al., 2007), they may instead reflect the covert mechanisms supporting evaluation.

A decade ago, a novel procedure named the Cue-Approach Training (CAT; Schonberg et al., 2014) was developed as a reliable means to change preferences without external reinforcements. The CAT procedure is a multiphase task consisting of an initial training session with individual items followed by a binary choice task. Initial studies using the CAT procedure showed that snack food preferences could be modified by associating images with a neutral cue and a rapid motor response (Bakkour et al., 2016, 2017; Schonberg et al., 2014). The task included a go/no-go training phase, during which a set of snack food stimuli was presented individually on the screen. A minority (∼30%) of the stimuli set was associated with a go cue, to which participants were required to respond with a rapid button press response (go stimuli). The remaining larger set of snack-food stimuli was passively presented for the same exposure period, without a cue or response (no-go stimuli). The training phase included several repetitions (runs) in which participants could learn the association of individual go stimuli with the cue and the response. The subsequent probe phase included binary choices between two snacks with similar initial subjective value for actual consumption (Bakkour et al., 2016, 2017; Schonberg et al., 2014). Preference change was demonstrated by enhanced preference for the go stimuli over no-go stimuli in the probe phase.

Dozens of studies demonstrated the efficacy of the CAT procedure in modifying preferences for a wide range of stimuli beyond snack foods, including healthy food items, faces, positive affective stimuli, and fractal art images (Aridan et al., 2019; Bakkour et al., 2016, 2017; Botvinik-Nezer et al., 2021; Chen et al., 2021; Salomon et al., 2018, 2019; Schonberg et al., 2014; Veling, Chen, et al., 2017; Zoltak et al., 2018). Several studies examined the long-term maintenance of the CAT effect in additional follow-up sessions and found that the preference modification effect persisted for months without any additional training sessions (Botvinik-Nezer et al., 2020; Chen et al., 2021; Salomon et al., 2018, 2019; Schonberg et al., 2014). The observation of preference modification following training without external reinforcement or feedback, and using individually presented items, suggests that CAT operates via a nonreinforced valuation pathway, potentially involving attention and motor neural circuits (Schonberg & Katz, 2020).

These robust results leave open the puzzle of why such training is effective: that is, what about it impacts later value-based choice. Several mechanisms have been proposed to explain nonreinforced learning paradigms. One prominent framework, preferences-as-memory, proposes that value can be conceptualized as retrieving of knowledge about alternatives from memory (Weber & Johnson, 2009). Relatedly, studies showed that go stimuli that were better remembered were also more preferred and that memory was positively correlated with likelihood of choosing go stimuli (Botvinik-Nezer et al., 2021; Chen et al., 2021). Attention has also been shown to play a key role in value-based decision making both on its own and in synergy with memory (Kraemer et al., 2022; Weilbächer et al., 2021). Eye-tracking studies of CAT provided support for the role of attention in the expression of modified preferences and found that during the probe task, participants spent more time viewing chosen Go stimuli (Bakkour et al., 2016; Salomon et al., 2019; Schonberg et al., 2014; Zoltak et al., 2020). This increased gaze time may indicate an attentional evidence-gathering process contributing to the higher likelihood of choosing go stimuli over no-go stimuli (Krajbich, 2019; Krajbich et al., 2010; Krajbich & Rangel, 2011). However, most of these studies did not find an enhanced gaze toward go stimuli when they were not chosen (Bakkour et al., 2016; Salomon et al., 2019; Zoltak et al., 2020), suggesting that attention alone may not fully account for the CAT preference modification effect.

Importantly, these theoretical explanations and their comparison to data in the task—primarily focused on the probe phase, where preferences were explicitly expressed, rather than on the training phase, where preferences are modified. The current article recenters on investigating the latter, in order to investigate what precisely during the training phase impacts later preference. Previous results provide some initial clues. First, the inclusion of no-go stimuli provides a control for mere exposure during training: Since all stimuli are presented individually for the same duration during training, mere

data curation, formal analysis, methodology, visualization, writing–original draft, and writing–review and editing. Nathaniel D. Daw played a supporting role in formal analysis, methodology, writing–original draft, and writing–review and editing and an equal role in supervision. Tom Schonberg played a lead role in funding acquisition, resources, and supervision, a supporting role in validation, writing–original draft, and writing–review and editing, and an equal role in conceptualization and methodology.

Correspondence concerning this article should be addressed to Tom Schonberg, School of Neurobiology, Biochemistry and Biophysics, Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 6997801, Israel. Email: Schonberg@tauex.tau.ac.il

exposure or viewing time alone cannot explain the preference modification effect. Several follow-up experiments manipulating training conditions collectively indicate the need for a speeded response toward the cue (Bakkour et al., 2016). In particular, no preference modification effects were found when the cue was presented simultaneously with the go stimulus, allowing participants a full second to respond, or when all go stimuli were presented in a block without interleaved no-go stimuli. These findings suggested that simple attention or motor response is insufficient to drive preference modification in the task. Instead, the rapid motor response during CAT and the motor preparation learning are crucial factors for nonreinforced preference modification with CAT.

However, it is still unclear what are the mechanisms operating during training at the individual item level where only a cue and speeded motor response in the absence of external reinforcement induce a preference change in the subsequent choice phase. A recent imaging study with CAT (Salomon et al., 2019) provided insight into the potential mechanisms at play during training. This study found that individual differences in the preference modification effect (measured during the probe phase) were correlated with increased neural activity in the supplementary motor area and striatum during the training phase. These regions were previously associated with motor planning (Ridderinkhof et al., 2004; Tanji, 1996) and reinforcement learning, respectively (Clithero & Rangel, 2014; Kable & Glimcher, 2009; O'Doherty et al., 2004; Pessiglione et al., 2006). The striatum has been suggested to be a key neural hub for learning due to its dual role in both reward processing and motor control (Collins & Frank, 2014; E. M. Tricomi et al., 2004). The striatal activity observed in relation to nonreinforced training and subsequent preference changes might reflect a process of internal reinforcement (Chew et al., 2021) during CAT, especially given the lack of external feedback. Thus, this neuroimaging results hint at the potential role of motor planning at the training phase linked to subsequent binary preference change.

These results are consistent with the hypothesis that preference modification rests in learning a specific sensorimotor association between individual stimuli and the go (vs. no-go) response. This might, in turn, give hints as to why the training increases later choice of that stimulus: past studies show that the brain tends to be biased toward go or active approach responses for rewards (e.g., Niv et al., 2007), and thus perhaps also is biased to treat stimuli associated with go as valuable. However, although this interpretation is plausible and broadly consistent with the framing of earlier CAT studies, it is important to stress that such a specific relationship has not yet been directly investigated or affirmatively demonstrated. In fact, it is not even demonstrated that CAT training reliably produces such associations: that is, anticipatory, stimulus-specific responses to the go cue. It remains possible, for instance, that some other distinction between go and no-go stimuli might drive the later difference in choices, for example, that differential attention to the former might mediate a difference in mere exposure. The present study tests the hypothesis that the specific stimulus–go association is associated with preference modification, rather than some other aspect of the stimuli. In particular, leveraging data from dozens of behavioral replications of this task in multiple laboratories around the world, in the present study, we aimed to identify a behavioral marker during the training that involves the training of individual items and a speeded button press response. We seek to investigate valuation mechanisms at the individual item level. We aimed to leverage the unique structure of CAT whereby training is conducted on individual items, while preference change is assessed in a subsequent binary choice phase. Therefore, we hypothesized that individual differences in the only active part during learning: button presses, could explain the variability in the subsequent preference modification phase observed after training with CAT.

To test our hypothesis, we developed a Bayesian computational approach to model an individualized learning marker derived from training motor response data at the individual item level. We then examined the association of this marker with the magnitude of preference modification observed in the subsequent binary choice probe phase. To develop this new computational framework, we utilized data from 28 experiments that employed the CAT procedure in the last decade. Then based on these modeling findings we designed and collected two new samples with a nonreinforced training procedure that directly manipulated the proposed learning mechanism. We aimed to close a full circle and establish a causal link between our model and preference change. This new modeling approach and procedure aimed to provide a framework for understanding the intrinsic mechanisms of nonreinforced preference modification through motor learning.

Thus, the current work consists of two main parts. In the first part, we developed a novel computational marker for learning, based on the only active component that varies during training: the button press. To achieve this, we used a unique data set based on a standardized nonreinforced preference modification task from 840 participants collected in 28 previous CAT experiments. We developed a computational model of reaction time (RT) patterns using a Bayesian framework based on an exploratory analysis of RT patterns during the training phase of these participants. The Bayesian model included an individualized parameter that captured the transition from cue-dependent to anticipatory responses, serving as a predictive marker of individual learning. We examined the association of the individualized learning marker with individual differences in preference modification. We further examined the hypothesis that nonreinforced preference modification operates both at the participant level (i.e., some participants learn better than others) as well as at a more granular item level (i.e., within participants, some stimuli are better learned than other).

In the second part of this work, we used the findings from our Bayesian model to develop a novel task design that allowed us to directly manipulate the computational learning marker. This manipulation aimed to demonstrate a causal relationship between the model and preference change. In the new design, we used two different cue-contingency conditions: 50% and 100% to manipulate learning difficulty. We hypothesized that the cue contingency manipulation would affect the motor response, that it would be reflected by the individualized learning marker devised in the first part, and most importantly predict subsequent preference modification in the probe phase. The second part included one preliminary experiment ($n = 20$), which was followed by a larger preregistered replication study ($n = 59$). All the hypotheses, experimental design and analyses plans were preregistered before data collection of the replication study began (https://osf.io/nwr4v).

The current work offers a unique investigation that offers novelty at the theoretical level by identifying an individualized computational marker for learning in CAT, providing empirical evidence that nonexternally reinforced preference modification occurs at the individual item level via motor learning cognitive mechanisms. The work further provides a rigorous methodological advancement by establishing a reliable quantification of this individualized learning, without asking participants to overtly reveal their preference. A computational marker for learning holds the potential to passively monitor learning in real-time and support the development of novel closed loop interventions.

Transparency and Openness

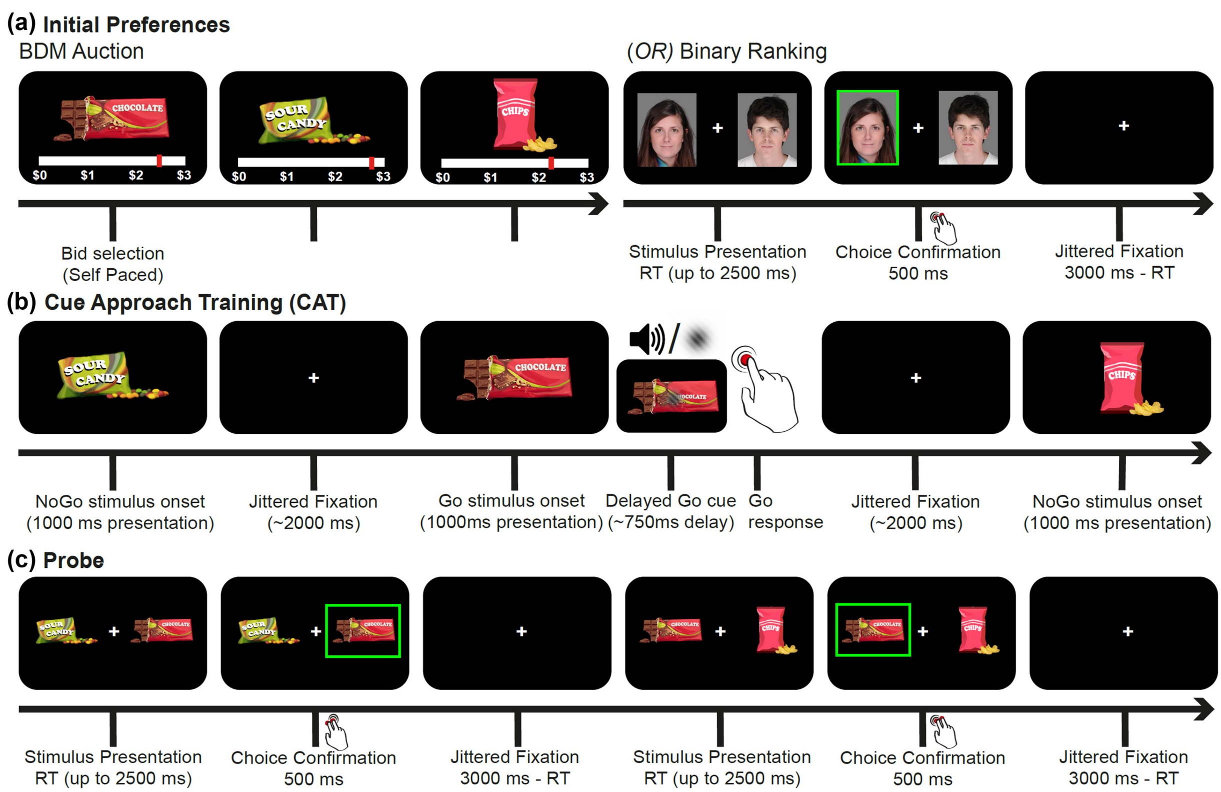
We preregistered the second part of our study in the Open Science Foundation at https://osf.io/nwr4v. All materials and analysis codes are available online at https://github.com/tomsalomon/CAT_Individualized_Learning. Preregistration hypotheses, experimental design data, and codes used for power analysis are available in the Open Science Foundation depository (https://osf.io/nwr4v).

Study 1: Meta-Analysis of CAT Studies

In the first part of the current work, we performed a meta-analysis of 28 CAT experiments with a total of $N = 860$ participants, which had been conducted by our research group and colleagues (see methods for sample size, publication origin, and demographics of the individual experiments). All experiments included the three main phases of the CAT procedure: initial preference evaluation, CAT, and a preference modification probe task (Figure 1; see detailed description in the Methods section). In the first phase, baseline preferences for a set of stimuli were evaluated using either a Becker-DeGroot-Marschak (BDM) auction procedure (Becker et al., 1964) for snack food stimuli, or a forced choice task (Salomon et al., 2018) for other stimulus types. Following the initial preference evaluation, participants completed the CAT phase. In this phase, approximately 30% of the stimuli (go stimuli) were consistently paired with a neutral go cue, prompting participants to respond with a rapid button press. The remaining stimuli (nogo stimuli) were presented passively without a cue or required response. Participants were instructed to wait for the go cue and respond as quickly as possible upon perceiving it. In the final probe phase, preference modification following CAT was evaluated. Participants were asked to indicate their preferred stimulus out of pairs of stimuli with similar initial value, in which one of the two stimuli was previously

Figure 1

General Outline of the Three Main Procedural Components of the Cue-Approach Training (CAT) Paradigm



Note. (a) Initial preference evaluation task. Baseline preferences for all stimuli were evaluated either with using a Becker–DeGroot–Marschak (BDM) auction (for consumable stimuli) or using a forced choice binary ranking task (for nonconsumable stimuli). (b) In the cue approach training (CAT) task, approximately 30% of stimuli were presented individually in association with a delayed cue (auditory or visual) to which participants responded with a rapid button press (go stimuli). All other stimuli were presented without the cue and response (no-go stimuli). (c) In the probe phase, preference modification was evaluated using a binary forced choice between pairs of stimuli of similar initial value, where one was a go stimulus and the other a no-go stimulus. Face images are from the copyright holder "Detecting Siblings in Image Pairs," by T. F. Vieira, A. Bottino, A. Laurentini, and M. De Simone, 2014, *The Visual Computer*, *30*(12), pp. 1333–1345 (https://doi.org/10.1007/s00371-013-0884-3). Copyright 2014 by Springer. Adapted with permission. RT = reaction time; *OR* = odds ratio. See the online article for the color version of this figure.

(during the training phase) a go stimulus, and the other was a no-go stimulus.

The meta-analysis study focused on identifying response patterns in the training task, which are the only active component that could indicate learning efficacy and predict the behavioral change measured in the subsequent probe task. Since the only variable component in the learning phase is RT, we hypothesized that changes in RT could be linked to learning. Due to the structure of the task including a staircase, RT alone could not be a predictor.

Method

Data Collection and Sample Sizes

For study 1 meta-analysis, the data from 28 previous CAT experiment were combined. The data set included 21 experiments from published works by our research group and collaborators (Aridan et al., 2019; Bakkour et al., 2016, 2017; Botvinik-Nezer et al., 2020, 2021; Salomon et al., 2018, 2019; Schonberg et al., 2014), as well as seven additional unpublished studies, which were collected as part of the preparations for previously published manuscripts, or manuscripts that are planned to be published in the future. The experiments included a median sample size of $n = 26$ ($M_{\text{sample size}} = 30$, range = [23, 70]; see Table 1 for information of the sample size in each experiment and the supplementary data for more detailed description of the unpublished data).

All participants gave their informed consent to take part in the experiments. In most experiments, participants received monetary compensation for their time, and in a few experiments, some participants took part in the experiment in exchange for course credit. All experiments were approved by the ethical review board of the institutes where they were performed (Tel Aviv University, The Hebrew University of Jerusalem, University of Texas at Austin, and McGill University).

Stimuli

The different experiments included in Study 1 examined the effect of CAT on preferences for various stimuli (see Table 1 for a summary of all experiments). In most of the experiments, the stimuli set comprised images of familiar local snack food items, popular in the United States (eight experiments) or in Israel (seven experiments), which participants received for actual consumption as part of the experiment. Other experiments used face stimuli of unfamiliar figures posing neutral expression from the Siblings data set (Vieira et al., 2014; two experiments), unfamiliar faces with neutral and happy expression from the Karolinska directed emotional faces data set (Lundqvist et al., 1998) (two experiments) or familiar faces of famous Israeli politicians (two experiments). Unfamiliar abstract stimuli of fractal art (adapted from https://fantasticfractals.com) were used in three experiments. Two experiments included positive affective stimuli from the international affective picture system (IAPS) data

Table 1
Experiment Included in the Meta-Analysis

Experiment	Stimuli	*n*	Training runs	Go cue	Publication (experiment number in publication[a])
1	Fractal art	25	12	Auditory	Salomon et al. (2018) (2)
2	IAPS positive	27	12	Auditory	Salomon et al. (2018) (3)
3	IAPS negative	28	12	Auditory	Salomon et al. (2018) (4)
4	Snacks (IL)	25	20	Visual	Salomon et al. (2018) (5)
5	Snacks (IL)	25	20	Neg. auditory	Salomon et al. (2018) (6)
6	Faces	25	20	Auditory	Salomon et al. (2018) 7)
7	Fractals	25	20	Auditory	Salomon et al. (2018) (8)
8	IAPS positive	29	20	Visual	Salomon et al. (2018) (9)
9	IAPS negative	32	20	Visual	Salomon et al. (2018) (10)
10	Faces: politic.	25	20	Auditory	Unpublished
11	Faces: politic.	39	20	Auditory	Unpublished
12	Faces	42	16	Auditory	Salomon et al. (2019)
13	Faces: affective	42	20	Auditory	Unpublished
14	Faces: affective	70	20	Auditory	Unpublished
15	Fractals	29	16	Auditory	Aridan et al. (2019)
16	Snacks (U.S.)	26	12	Visual	Unpublished
17	Snacks (IL)	30	16	Auditory	Botvinik-Nezer et al. (2021) (1)
18	Snacks (IL)	25	16	Auditory	Botvinik-Nezer et al. (2021) (Pilot)
19	Snacks (IL)	23	12	Auditory	Unpublished
20	Snacks (IL)	25	20	Auditory	Unpublished
21	Snacks (U.S.)	29	12	Auditory	Schonberg et al. (2014) (1)
22	Snacks (U.S.)	25	8	Auditory	Schonberg et al. (2014) (2)
23	Snacks (U.S.)	25	12	Auditory	Schonberg et al. (2014) (3)
24	Snacks (U.S.)	27	16	Auditory	Schonberg et al. (2014) (4)
25	Snacks (U.S.)	26	16	Auditory	Schonberg et al. (2014) (7)
26	Snacks (U.S.)	30	12	Auditory	Bakkour et al. (2017)
27	Snacks (U.S.)	25	16	Auditory	Bakkour et al. (2016)
28	Snacks (IL)	36	16	Auditory	Botvinik-Nezer et al. (2020)

Note. IAPS = international affective picture system; IL = Israel; U.S. = United States.
[a] In articles with multiple experiments, we note in parenthesis the experiment number or title as can be found in the related article.

set, and two experiments included negative affective stimuli from the IAPS data set (Lang et al., 2008).

The face and snack stimuli were modified using Photoshop to remove any background features and create visual standardization within each experiment. In each experiment, all stimuli were colored images of identical dimensions, where the subject of the image (the snack or the face) was centered in the middle of the image frame, and the background was replaced with homogeneous background (either black or gray). In experiments with fractal art and IAPS stimuli, the images were only cropped to identical pixel dimensions.

Go Stimuli Used in the CAT Task. In most experiments (23 of 28), a neutral auditory cue was used as go cue during the CAT task. In the remaining experiments, a neutral visual cue (Experiments 4, 8, 9, and 16) or an aversive auditory cue (Experiment 5) was used. The visual cue included a semitransparent Gabor shape, presented on top of the associated go stimuli.

Procedure

While the different experiments diverged in several aspects of their procedure, all experiments maintained three key phases: an initial-preferences evaluation phase, followed by a training phase, and a probe phase (see Figure 1 for illustration of the procedural design).

Initial Preferences Evaluation Phase. In the first phase of the experimental procedure, participants were exposed to the complete set of stimuli for the first time and were required to indicate their subjective preferences using one of two tasks. To evaluate participants' preferences in experiments using consumable snack-food stimuli, participants performed a Becker–DeGroot–Marschak (BDM) auction procedure (Becker et al., 1964). Participants were allocated with either US$3 (in experiments conducted in the United States) or 10 ILS (in experiments conducted in Israel; approximately equivalent to US$3.1) which were used to bid on the snack-food stimuli, presented one by one. Participants were informed that at the end of the experiment, one of the trials will be randomly selected, for which the computer will generate a counter bid. If the participants' bid was higher than that of the computer, they were required to purchase the snack for the lower price bided by the computer. Participants purchased the snack for actual consumption at the end of the experiment. However, if the computer's bid was higher, participants were not allowed to purchase the items and kept the allocated sum of money. Participants were explicitly instructed that the best strategy for the task is to indicate their true subjective preference. Prior to their participation, participants were asked to fast for at least 3 hr, to make sure they were hungry and incentivized to purchase the food items, according to their subjective preferences.

In experiments with nonconsumable stimuli (such as fractals and faces), which are less appropriate to be evaluated using a monetary scale, preferences were evaluated using a binary ranking procedure. Participants were presented on each trial a random pair of stimuli and were required to choose the stimuli they preferred better. Based on the idea of choice transitivity, binary choices were quantified to produce subjective value ranks using the Colley Matrix ranking procedure (Colley, 2002).

Following the BDM or binary choice task, stimuli were ranked-ordered according to subjective preferences. The ranks of the stimuli were used as a basis to form two value groups—one of high-value stimuli (above-median rank) and a group of low-value stimuli (below median rank). The size of the value groups differed between experiments (having 8, 12, or 16 stimuli per value group). In each of the value groups, half of the stimuli were allocated to be associate with the go stimuli and response in the subsequent CAT task (go stimuli), and the other half of stimuli was allocated to be presented in the training phase of without a go cue (no-go stimuli).

Allocation for go and no-go stimuli within each value category maintained an equal mean rank for go and no-go stimuli—for example, for a high-value group consisting of eight stimuli, the stimuli ranked 8, 11, 12, and 15 were allocated to be go stimuli, while stimuli ranked 9, 10, 13, 14 were allocated to be no-go stimuli, such that both allocations were characterized with a mean rank = 11.5. The go/no-go allocation was counterbalanced across participants.

Cue Approach Training. In the CAT task, stimuli were presented individually on the screen center for a fixed duration of 1 s (except for Experiment 26, where the duration was extended to 1.2 s for compatibility with the functional imaging protocol). In each experiment, all training stimuli were presented once in each training run; thus, the number of training runs indicates the number of stimulus repetitions. A fixed proportion of stimuli per experiment (commonly ~30% of stimuli; range 25%–40%) were go stimuli. When a go stimulus appeared, a delayed go cue appeared after a go signal delay. Participants were asked to respond to the go cue with a button press as rapidly as possible, before the stimulus offset (1 s after the stimulus onset). The go signal delay was adjusted according to the participant's performance—a failure to respond before stimulus offset resulted in 50 ms shortening of the go signal delay (thus reducing the task difficulty for the following trial), while a successful response on time resulted in 16.667 ms increase of the next go signal delay (making the task more difficult; 1:3 ratio of signal delay increase to decrease). The go signal delay commonly started at 750 ms and ranged around 700 ms (across experiments, $M = 693.48$, $SD = 101.34$).

The training phase in each experiment, consisted of 8–20 training runs (see Table 1). Participants were not informed in advance of the contingency between go signal and go stimuli. However, as they were repeatedly exposed to the same stimuli, in the later runs of the task, some of the participants were able to identify the go stimulus and cue association, thus producing accurate faster responses, sometimes even preceding the go cue onset. This measurement of RT after or before appearance of go stimuli was used in this current work as our main behavioral measurement for the CAT learning model. The details of the model are discussed below.

Probe. In the final probe phase, preference modification following CAT was evaluated. In the probe task, participants were presented with pairs of stimuli of the similar initial subjective value (both high-value stimuli and both low-value stimuli). In each pair, one of the stimuli was a go stimulus, and the other was a no-go stimulus. Preference modification was evaluated as the proportion of trials in which participants chose the go stimulus over the no-go stimulus, above and beyond the expected 50% chance level. This measurement was used as the main outcome variable which indicated preference modification following CAT.

Analysis

In Study 1, we aimed to identify a marker for learning during the CAT task, whereby individual items appear, and participants need to press a button after a cue is presented. As pressing is the only variable action that could lead to subsequent preference changes we examined RT patterns in the task. In an exploratory analysis of the meta-analysis data, we identified that as training progressed, mean RT in the task was reduced. We thus hypothesized that as the training task progresses, with repeated presentations of the items, participants were able to identify and learn the stimulus–cue contingency pattern, and thus could generate faster go responses, which did not rely on the delayed go cue onset, but rather on the go stimulus onset itself. We formally modeled this process in a single experiment before testing it on the entire meta-analysis data set.

Computational Model of CAT. To model our proposed cognitive mechanism, we utilized a Bayesian modeling approach of the RT in the CAT task, using the R implementation of Stan programing language (Stan Development Team, 2020). RTs were modeled as a mixture of two Gaussian distributions (Equation 1)—one Gaussian of shorter mean RT, representing the early anticipatory responses generated when the stimulus–cue association is predicted by the participants; and a second Gaussian with a later mean RT following the cue onset (i.e., a distribution representing standard cue-dependent responses).

$$\begin{aligned}&\mathrm{RT}_{t,i} - \mathrm{Cue}_{t,i} = N(\mu_1, \sigma_{\varepsilon_1})(\theta_{t,i}) + N(\mu_2, \sigma_{\varepsilon_2})(1-\theta_{t,i})\\&\theta_{t,i} = \Phi(-3.1 + \theta_{\mathrm{slope}_i}\mathrm{Run}_{t,i})\\&\theta_{\mathrm{slope}_i} \sim N(\theta_{\mathrm{slope}}, \sigma_{\theta_{\mathrm{slope}}})\\&\mu_1 < 0, \mu_2 > 0; t\text{–trial index}; i\text{–participant index.}\end{aligned} \tag{1}$$

The $\theta_{t,i}$ mixture probability of the two Gaussian distribution was used to determine the proportion of trials; participants were expected to produce early anticipatory response (Equation 1). It was defined using a linear function of time-dependent $\theta_{\mathrm{slope}_i}$ parameter, which was multiplied by the scaled training run index (training repetition), scaled to 0–1, where 0 and 1 indicating the first and last (20th) training repetition, respectively. This $\theta_{\mathrm{slope}_i}$ parameter was fitted individually for every participant and was therefore defined conceptually as the individualized learning parameter of interest. To scale the linear function to a range suitable for proportions (0–1), we used the normal cumulative distribution function as a link function.

The baseline proportion of early anticipatory responses (at the first training run) was fixed at the value $\theta_0 = -3.1$, corresponding with probability of 0.1% to generate an anticipatory response at the very first run. Thus, the individualized learning parameter could be interpreted as the sole parameter affecting the mixture proportion for the two Gaussian distributions, for example, $\theta_{\mathrm{slope}_i} = 3.1$ would beinterpreted as a $\Phi(0) = 0.5$, 50% proportion of anticipatory responses at the final (20th) training run. The individually fitted $\theta_{\mathrm{slope}_i}$ parameter was determined in the current work as the computational marker of learning and was our main parameter of interest.

The choice to model participants' learning using a fixed intercept was made to consolidate all observed between-participant differences in learning into a single key parameter: the learning slope (θ_{slope}). This simplification allows for direct comparison between individuals based solely on their rate of transition to anticipatory responses. In other words, any variation in the proportion of anticipatory responses predicted by the model is entirely attributed to the θ_{slope} parameter. While using individualized intercepts could capture more nuanced patterns, it would introduce additional complexity and potentially hinder the interpretability of the model. We acknowledge that this is a simplifying assumption; however, it is supported by the notion that participants are unlikely to exhibit intentional anticipatory responses in the very first training run due to the lack of prior knowledge about cue–stimulus associations. To account for individual differences in θ_{slope}, we introduced $\sigma_{\theta_{\mathrm{slope}}}$, a parameter analogous to the random effects in mixed-model regression: $\sigma_{\mathrm{slope}_i} \sim N(\mu_{\theta_{\mathrm{slope}}}, \sigma_{\theta_{\mathrm{slpoe}}})$. See Supplemental Code S1 for a complete specification of the model parameters and priors.

The model's parameters were evaluated using Markov-chain Monte Carlo (MCMC) gradient algorithm, implemented with RStan (Stan Development Team, 2020). Each model was evaluated four independent times, with chains length of 2,000 (1,000 chain links burnout time). The mean values of each parameter of the converged model were used as the parameter estimates and are reported along with 95% credible interval (CI). All reported results converged onto stable solution with $\hat{R} = 1$.

To maintain stable interpretation of μ_1 and μ_2 as the centers of anticipatory and cue-dependent responses, respectively, an upper threshold of 150 ms from cue onset was imposed on the μ_1 parameter, and a lower threshold of 200 ms was for μ_2 parameter. Thus, in every chain, μ_1 was imposed the role of the earlier anticipatory responses. Alternative formulation forms of the computational model were also considered and tested, including using 0 ms as both upper and lower thresholds for the distribution means, using a log-normal distribution instead of a normal distribution, and fitting a unique μ_{1_i} parameter for each participant. However, these models either did not converge when applying to the entire data set or converged with some issues. Initially we used 0 ms as an upper threshold for μ_1, which caused the parameter to converge at this maximal threshold. Using a unique μ_{1_i} parameter for each participant was theoretically favorable (and indeed used in Study 2), but could not be applied in Study 1, potentially either due to computational challenges (fitting many more interdependent parameters) or due to the complex structure of the data (e.g., where anticipatory responses were intertwined with late responses due to changing go signal delay).

Computational Model of CAT With Stimulus-Level Parameters. In an exploratory model, designed following the analysis of Study 2, we aimed to fit a stimulus-level learning parameter. For each $\mathrm{stimulus}_s$ which was presented to $\mathrm{participant}_i$, we fitted an individualized $\theta_{\mathrm{slope}_{i,s}}$ parameter, which was used instead of the participant-level individualized learning parameter ($\theta_{\mathrm{slope}_i}$), see Supplemental Code S2. Initially, we attempted to model a participant-level dependence, by modeling $\theta_{\mathrm{slope}_{i,s}}$ as a parameter derived from higher level $\theta_{\mathrm{slope}_i}$ participant-level parameter, similarly to the design in Study 2 (see below). However, such model with tens of thousands of $\theta_{\mathrm{slope}_{i,s}}$ parameter estimates that covaried with hundreds of $\theta_{\mathrm{slope}_i}$ parameters was computationally too demanding. To simplify the model structure, $\theta_{\mathrm{slope}_{i,s}}$ parameter was modeled independently of participants' identity (i.e., not considering within-participant possible effect; see Supplemental Code S2).

Deviation From Previous Version and Preregistration. In our preregistration, we used an identical Stan model with one key difference—to enforce an order in which μ_1,σ_1 would relate to early anticipatory RTs and μ_2,σ_2 would relate to late cue-dependent RTs, a different restriction was set in which the upper and lower limit of μ_1 and μ_2, respectively, were set to 0 (cue onset), instead of the final limits set to 150 and 200 ms. Running the model with these different restrictions converged to a stable solution. Early anticipatory responses were modeled as having an earlier mean ($\mu_1 = -0.04$ ms, 95% CI [−0.15, 0.00], $\sigma_{\varepsilon_1} = 286.00$, 95% CI [285.55, 286.43]) and late cue-dependent responses were very similar to the new model

(μ_2 = 282.34 ms, 95% CI [279.73, 284.94], σ_{ε_1} = 75.91, 95% CI [75.56, 76.25]). Other parameters were also very similar (θ_{slope} = 3.48, 95% CI [3.19, 3.77]), with variation between participants ($\sigma_{\theta_{\text{slope}}}$ = 4.11, 95% CI [3.87, 4.35]). All reported association with probe phase remained very similar.

While the model showed no formal convergence errors (see trace plots of parameters in Supplemental Figure S9), the μ_1 parameter estimate seemed to have been affected by the artificial limitation which was imposed on its maximal value for technical reasons (to avoid inversion with μ_2 across the different chains). Throughout the different chains, the final parameter converged around the upper limit set to 0 (the Cue onset). After several experimental attempts to change this limit, we found the increasing μ_1 lower limit to 150 ms resulted in a better solution, in which μ_1 parameter estimate did not converge around the uppermost limit. This model is the final model we chose to use and report here.

Probe Analysis. To evaluate the effect of CAT on preferences, we analyzed the proportion of trials in which participants chose the go stimulus over the no-go stimulus in the probe phase, using mixed-model logistic regression. As go and no-go stimuli were matched based on initial value, under the null hypothesis, participants were expected to choose go stimuli at 50% of trials (log-odds = 0; odds = 1). The results of this analysis were of main interest in previous publications and are reported in the current work for unpublished data (see Supplemental Materials).

While early work with CAT task showed a differential effect of value on CAT effect on preferences—that is, CAT usually had induced more prominent preference modification for stimuli of initial high value, compared to stimuli of initial low value (Bakkour et al., 2016; Schonberg et al., 2014), more recent work with CAT found that this effect was not a dominant feature of CAT (Botvinik-Nezer et al., 2020, 2021; Salomon et al., 2018, 2019). Thus, in the current work, we pooled together data of stimuli with both high- and low-initial value.

Probe and Individualized Learning Parameter Association. Our main analysis of interest aimed to evaluate the association of the CAT preference modification effect with the individualized learning parameter θ_{slope_i}, derived from the computational modeling of CAT. In a mixed-model logistic regression analysis implemented with lme4 R package (Bates et al., 2015), the number of trials each participant chose the Go versus no-go stimuli was explained using the independent variable of θ_{slope_i} of each participant, as calculated in the preceding CAT task. The data were aggregated over all participants in the 28 experiments. To account for the aggregation of different experiments in the data set, we also included random intercept and random slope terms for each of the 28 experiments, see logistic regression formula in Equation (2).

$$\text{Choice} \sim 1 + \theta_{\text{slope}_i} + (1 + \theta_{\text{slope}_i}|\text{Exp.}). \tag{2}$$

In an additional exploratory analysis (which was conceived after the analysis of Study 2), we modeled choices using a stimulus-specific $\theta_{\text{slope}_{i,s}}$ parameter. We aimed to examine if modeling a learning parameter at a more precise stimulus-level (contrary to the preregistered participant-level θ_{slope_i} parameter) would serve as a more accurate predictor of future preference modification. For this aim, we merged the $\theta_{\text{slope}_{i,s}}$ which were estimated for each contingency with the probe data. Choices (number of trials each go stimulus$_s$ was chosen or not by each participant$_i$) were modeled with two additional models using $\theta_{\text{slope}_{i,s}}$ as an independent variable. In the first model, $\theta_{\text{slope}_{i,s}}$ parameter estimates were used instead of θ_{slope_i} estimates. Since $\theta_{\text{slope}_{i,s}}$ varied within participants, which were nested within experiment, a nested random slope effect for $\theta_{\text{slope}_{i,s}}$ was also added (Equation 3). In two additional models, probe choices were explained using a full model with both $\theta_{\text{slope}_{i,s}}$ and θ_{slope_i} (Equation 4) and a nested model with the same independent variables except for the fixed and random effects of $\theta_{\text{slope}_{i,s}}$ (Equation 5). Using a likelihood ratio test for nested models (comparing model Equation 4 with the nested model of Equation 5), we examined whether $\theta_{\text{slope}_{i,s}}$ provided additional explanatory above and beyond θ_{slope_i}.

$$\text{Choice} \sim 1 + \theta_{\text{slope}_{i,s}} + (1 + \theta_{\text{slope}_{i,s}}|\text{Exp./Participant}), \tag{3}$$

$$\begin{aligned}\text{Choice} \sim 1 + \theta_{\text{slope}_{i,s}} + \theta_{\text{slope}_i} + (1 + \theta_{\text{slope}_{i,s}}|\text{Exp./Part.})\\ + (\theta_{\text{slope}_i}|\text{Exp.}),\end{aligned} \tag{4}$$

$$\text{Choice} \sim 1 + \theta_{\text{slope}_i} + (1 + \theta_{\text{slope}_i}|\text{Exp.}). \tag{5}$$

Effect Size Estimation for Mixed Models With R^2_{GLMM}. For the reported logistic regression mixed model, we included an additional effect size score using a generalized linear mixed effect model (GLMM) R^2 estimate, based on the work by Nakagawa and Schielzeth (2013), implemented with R's Multi-Model Inference (MuMIn) package (Bartoń, 2020; Johnson, 2014; Nakagawa et al., 2017; Nakagawa & Schielzeth, 2013). Like R^2 in linear mixed model, R^2_{GLMM} was used to quantify the relative proportion of variance accounted by the three generalized mixed-model's variance components: variance explained by the fixed effects (σ^2_f), variance explained by the random effects (σ^2_α), and unexplained residual variance (σ^2_ε). In accordance with the developers' suggestion (Nakagawa et al., 2017; Nakagawa & Schielzeth, 2013), we report here two R^2 scores—the marginal R^2 (Equation 6), which signifies the proportion of total variance accounted by the fixed effects (σ^2_f); and the conditional R^2 (Equation 7), which signifies the relative proportion of variance explained by both the fixed and random effects of the model.

$$R^2_{\text{GLMM}(m)} = \frac{\sigma^2_f}{\sigma^2_f + \sigma^2_\alpha + \sigma^2_\varepsilon}, \tag{6}$$

$$R^2_{\text{GLMM}(c)} = \frac{\sigma^2_f + \sigma^2_\alpha}{\sigma^2_f + \sigma^2_\alpha + \sigma^2_\varepsilon}. \tag{7}$$

In mixed models examining the contribution of stimulus-level computational marker ($\theta_{\text{slope}_{i,s}}$; Equations 3–5), R^2_{GLMM} values were similarly evaluated. However, it is important to note that the basic unit of analysis in this model (choices per go stimulus within participant) differs from the unit of analysis in Equation 2 (choices per participant). Thus, the R^2_{GLMM} values are not comparable between the two models. The $R^2_{\text{GLMM(m)}}$ used the participant-level model is the most similar in structure and interpretation to "traditional" R^2 in ordinary linear model, representing the proportion of variance explained by the (fixed) effects of interest, using the participant as the basic unit of analysis.

Results

RT Analysis of CAT

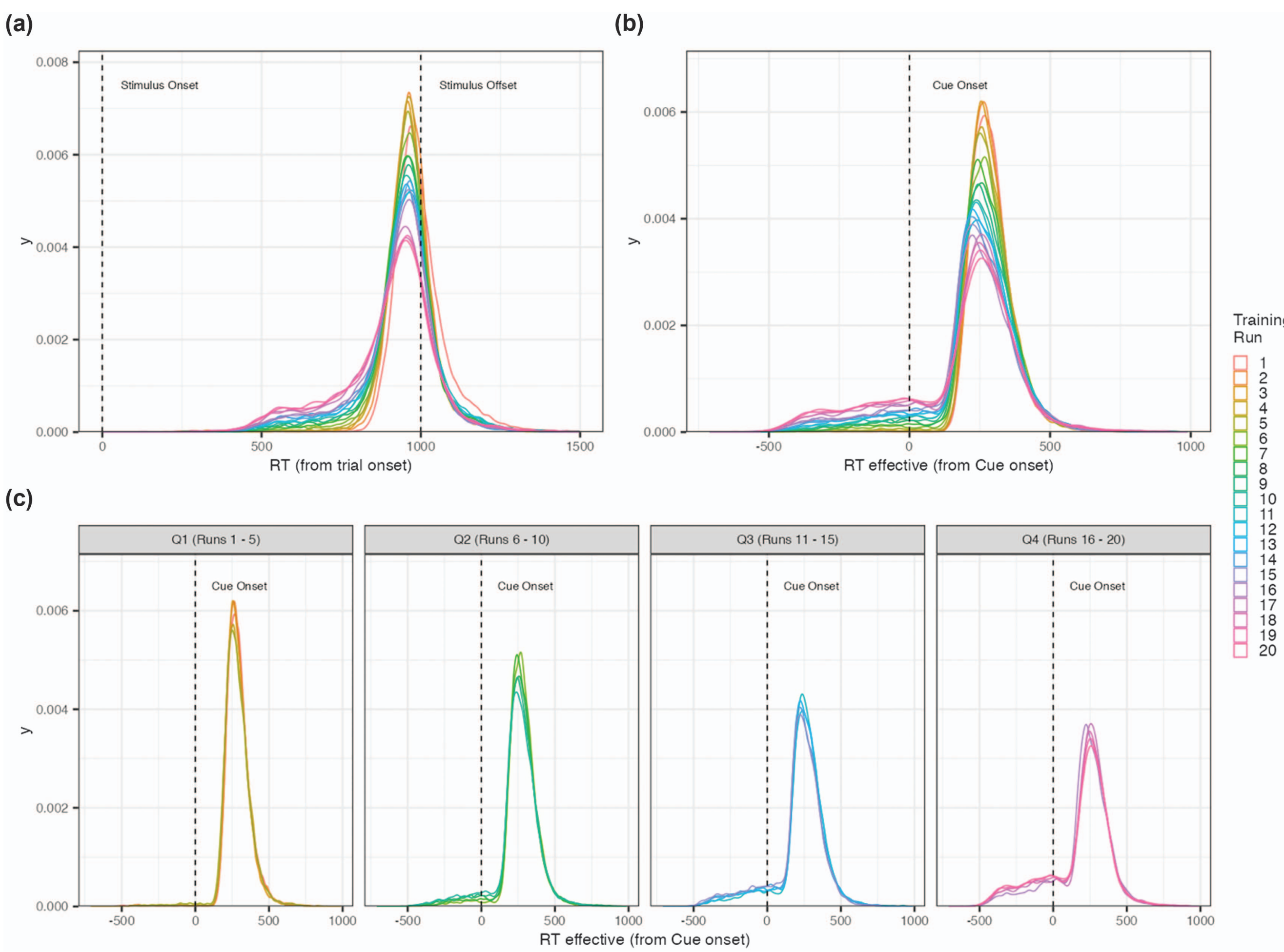
The training task in the different experiments consisted of 8–20 training runs (the number of training runs varied between experiments and identical within each experiment). As a reminder, during each run, all go and no-go stimuli were presented once individually on-screen. During no-go trials, stimuli were presented without a go cue and required no response. Whenever go stimulus appeared (approximately 30% of trials), a go cue was followed, to which participants were asked to respond with a rapid button press, before stimuli offset. Participants were instructed to respond *only* when they hear or see the cue. Each trial started with stimulus onset, which was presented on screen for 1,000 ms. During go trials, a go cue appeared approximately 750 ms following stimulus onset (thus leaving participants approximately 250 ms to respond). The cue onset changed according to participants' performance, so that each successful go response was followed with an increase in go onset time (thus leaving less time to respond and making the next trial more challenging; see detailed description in the methods section).

To identify unique response patterns, we performed an exploratory analysis of the RT distribution during go trials. Before analyzing RT in the CAT task, invalid trials were excluded. In total, 1,681 trials out of 182,464 go trials were excluded (0.92% of trials), due to participants failing to respond within the allocated 1,500 ms from stimulus onset time frame (0.91% of trials), and trials in which the response was shorter than a threshold of 100 ms (0.01% of trials), as these trials are likely to be indicative of inattentional response.

Examining the RT density distribution in the training task as a function of training run, revealed a distinct pattern, wherein as training progressed, participants' RTs were less homogeneous and started to form a growing peak of early RTs (Figure 2a). In all 28 experiments, the go cue changed on each trial, according to participants' performance (see methods), thus, a more informative

Figure 2

Reaction Time Distribution of Data Pooled From Meta-Analysis With 28 Studies



Note. Density plots of (a) RTs time-locked to the trial onset and (b) RTs time-locked to the varying go signal onset (*effective RT*). Color indicates training run (repetition). As training progressed, the RT distribution shifted from a cue-dependent unimodal distribution to left-tailed distribution with increasing proportion of early anticipatory responses. (c) A clear differentiation in effective RT is apparent when splitting the data (from Figure 2b) to four quartiles, according to the training run (each quartile indicating five consecutive training runs). RT = reaction time.

measurement of RT was the time from cue onset (RT minus cue onset), referred to here as *effective RT*. In earlier training runs, when the go stimuli were associated for the first time with the go cue, most effective RTs were clustered around a unimodal center, approximately 300 ms following the go signal onset (effective RT $M = 290.52$ ms) with 97.5% of effective RTs larger than 175.69 ms and 99% of effective RTs larger than 144 ms. Thus, we use here an empirical quantile of 145 ms as threshold to define rapid *anticipatory responses*, as these responses can be elicited when the participants predicted a go cue would appear. As training progressed, a consistent pattern appeared in the effective RT data—the main peak of the RT distribution was reduced, while a growing portion of participants' responses consisted of faster anticipatory RTs, in many cases preceding the go cue onset (Figure 2b). For example, while the proportion of anticipatory responses comprised only 1% of RTs in the first training run, they comprised 2.7% of RTs by the 5th run, 12.1% of the trials by the 10th run, 20.1% of trials by the 15th run, and finally 28.9% of trials in the final 20th run (aggregated across all samples).

Thus, RT data suggest a temporal-dependent response pattern. As training progressed, participants tended to rely less on the go cue (late cue-dependent RTs) and generated more early anticipatory responses. Interestingly, the mode of the RT distribution seemed to remain stable, and only the relative proportion of the two distributions changed. We hypothesized that this transition pattern from the cue-dependent response to a stimulus-triggered anticipatory response reflected a process of learning about the stimuli that could be associated with behavioral change in the subsequent preference probe. Thus, we next sought to develop a model to characterize the strength of this effect within individuals.

Individualized Learning Computational Model. Based on the unique RT pattern observed in the exploratory analysis of CAT task data, we developed a novel computational Bayesian model. To capture the form of RT distribution with rapid anticipatory responses and slower cue-dependent responses, we modeled the effective RT as a mixture of two Gaussian distributions, with two different means (μ_1 and μ_2 free parameters), two standard deviations (σ_{ε_1} and σ_{ε_2}), and a mixture proportion ($\Theta_{i,t}$) indicating the relative mixture proportion for participant$_i$ at trial$_t$. The model was implemented with the Stan probabilistic programming language (Stan Development Team, 2020), which optimized the model parameter estimates using a Markov-chain Monte Carlo (MCMC) approach. See the Methods section for a detailed description of the model and Supplemental Code S1. The first Gaussian was restricted to have a lower mean than the second Gaussian to maintain their correspondence across different MCMC sampling chains (i.e., $\mu_1, \sigma_{\varepsilon_1}$ would correspond with the early anticipatory RT distribution). To capture the gradual change in anticipatory responses over training, the mixture proportion was modeled as a time-dependent variable, expressed by the following formula:

$$\Theta_{i,t} = \phi(\theta_0 + \theta_{\mathrm{slope}_i}\mathrm{Run}_{i,t}), \tag{8}$$

where the mixture probability $\Theta_{i,t}$ for participant$_i$ at trial$_t$ was modeled as a linear function (with a fixed θ_0 intercept and random participant-level $\theta_{\mathrm{slope}_i}$-free parameter) of training run (where the $\mathrm{Run}_{i,t}$ ranged between [0,1]). The linear trend was scaled nonlinearly to probability-appropriate values of [0, 1] range using a normal distribution cumulative distribution function (normal CDF; denoted with ϕ), resulting in a sigmoid-like link function.

To associate all learning with a single parameter and a simple interpretation of $\theta_{\mathrm{slope}_i}$-free parameter, we chose to fix the θ_0 intercept term, which represents the mixture proportion at the very first run ($\mathrm{Run}_{i,t} = 0$). The θ_0 intercept term was fixed at -3.1, corresponding with $\Theta_{i,t} = \phi(-3.1) \approx 0.1\%$ mixture proportion at the very first run for all participants. Thus, higher $\theta_{\mathrm{slope}_i}$ parameter estimate indicated an individual participant with faster transition from slow cue-dependent RTs to fast anticipatory responses, which corresponded with larger $\Theta_{i,t}$ mixture proportion at the very last run. For example, parameter estimate of $\theta_{\mathrm{slope}_i} = 3.1$, indicated that at the last run ($\mathrm{Run}_{i,t} = 1$), the mixture proportion of anticipatory responses would be $\phi(0) = 50\%$, while $\theta_{\mathrm{slope}_i} = 4.745$ would correspond with mixture proportion of $\phi(1.645) \approx 95\%$, at the last run. We hypothesized that this individualized parameter would be indicative of improved learning in the training phase of CAT and would also be positively associated with stronger preference modification effect, measured in the subsequent probe phase.

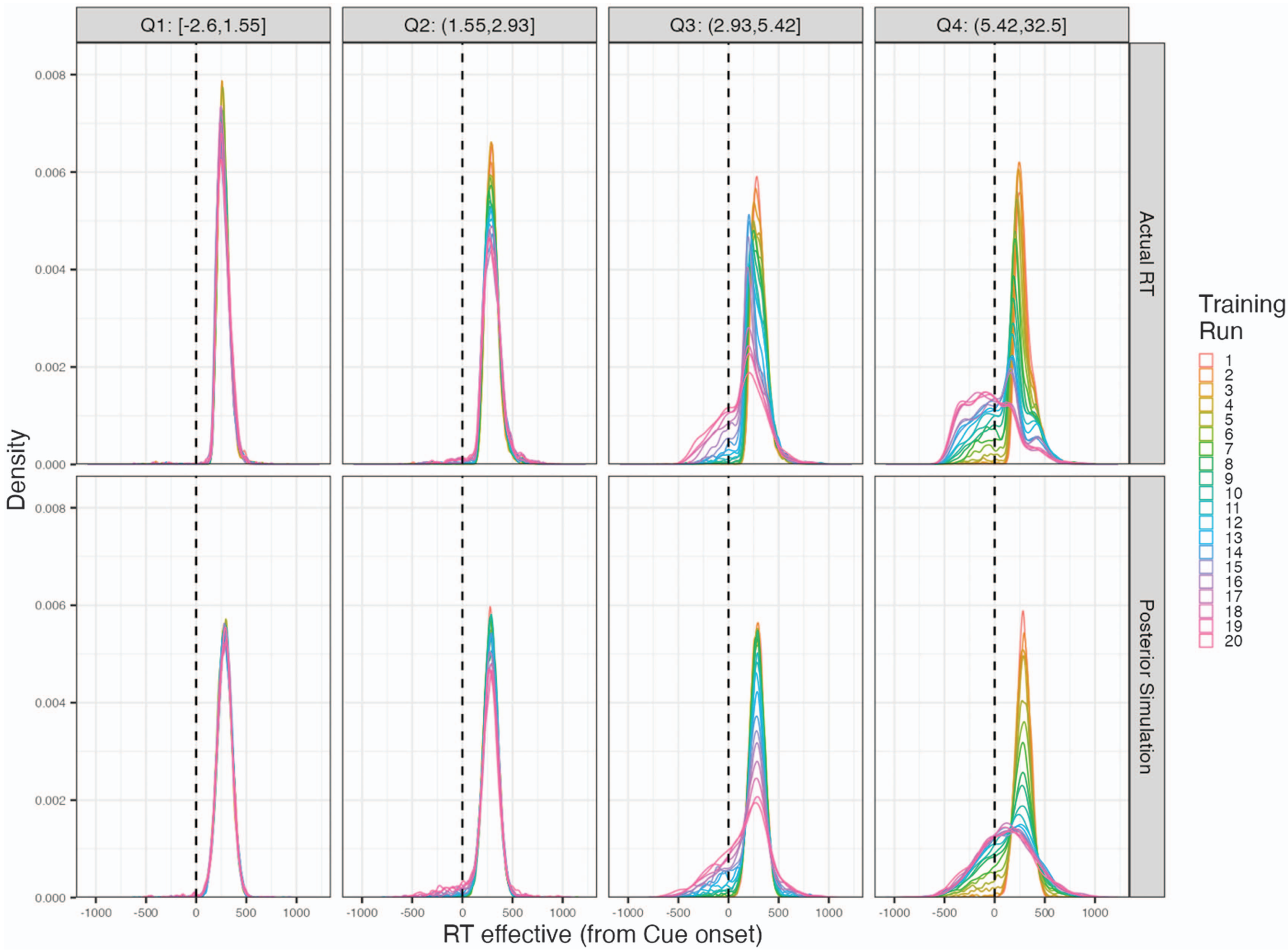
Fixing the expected anticipatory responses to a very low number in the first run was an assumption we chose to adapt in the current work to be able to quantify different participants' learning using $\theta_{\mathrm{slope}_i}$ as a single comparable parameter which captured all aspects related to learning. If θ_0 intercept term would also have been left as a free parameter, we would have needed to account for both parameters and their interaction to define a computational learning marker.

After running four independent chains, the Bayesian model converged to a stable solution (see Supplemental Figure S1). The converged model estimated effective RTs as a time-dependent mixture of two RT distributions: one of early anticipatory responses ($\mu_1 = 111.76$ ms, 95% CI [108.19, 115.41], $\sigma_{\varepsilon_1} = 275.00$ ms, 95% CI [272.72, 277.18]) and late cue-dependent responses ($\mu_2 = 284.15$ ms, 95% CI [283.69, 284.60], $\sigma_{\varepsilon_2} = 71.35$ ms, 95% CI [70.98, 71.72]). Overall, participants demonstrated an increase in proportion of anticipatory responses as training progressed, as manifested in the positive group-level parameter ($\theta_{\mathrm{slope}} = 4.24$, 95% CI [3.89, 4.60]), with variation between participants ($\sigma_{\theta_{\mathrm{slope}}} = 5.21$, 95% CI [4.90, 5.53]), indicating some participants made faster transitions and some generated nearly only cue-dependent responses (non-positive $\theta_{\mathrm{slope}_i}$ estimates).

Posterior predictive checks (simulated distributions of RT based on estimated model parameters) revealed a good fit of the model to the actual data. Simulated posterior distributions recreated the patterns observed empirically, showing more rapid transition to anticipatory responses in participants with higher $\theta_{\mathrm{slope}_i}$ parameter estimate (Figure 3; Supplemental Figure S2).

Learning Parameter Association With Choices. Previous findings with CAT consistently showed that preferences for go stimuli were enhanced following CAT, as manifested in choice behavior during the probe phase. When presented with a choice between a go stimulus and a no-go stimulus of similar initial value, participants consistently chose the go stimulus (Aridan et al., 2019; Bakkour et al., 2017; Botvinik-Nezer et al., 2020, 2021; Salomon et al., 2018, 2019; Schonberg et al., 2014; Veling, Chen, et al., 2017; Zoltak et al., 2018). To test the hypothesis that probe performance was associated with RT patterns during CAT, we examined whether variation in the slope parameter fit to RTs correlated with probe performance. Note that, under the model, the slope parameter controls the final proportion of

Figure 3
Actual RT Distributions Versus Simulated Posterior Distributions, by Quantile Group



Note. Participants were categorized into four equal quantile groups, according to their parameter estimates (denoted here as Q1–Q4; columns). Participants with higher parameter estimates were characterized with faster transition to anticipatory responses (top row). Posterior simulated RT distributions using mixture of Gaussians (bottom row) recreated relatively well this transition pattern. Vertical dashed line represents cue onset. See also Supplemental Figure S2 for more detailed comparison. RT = reaction time.

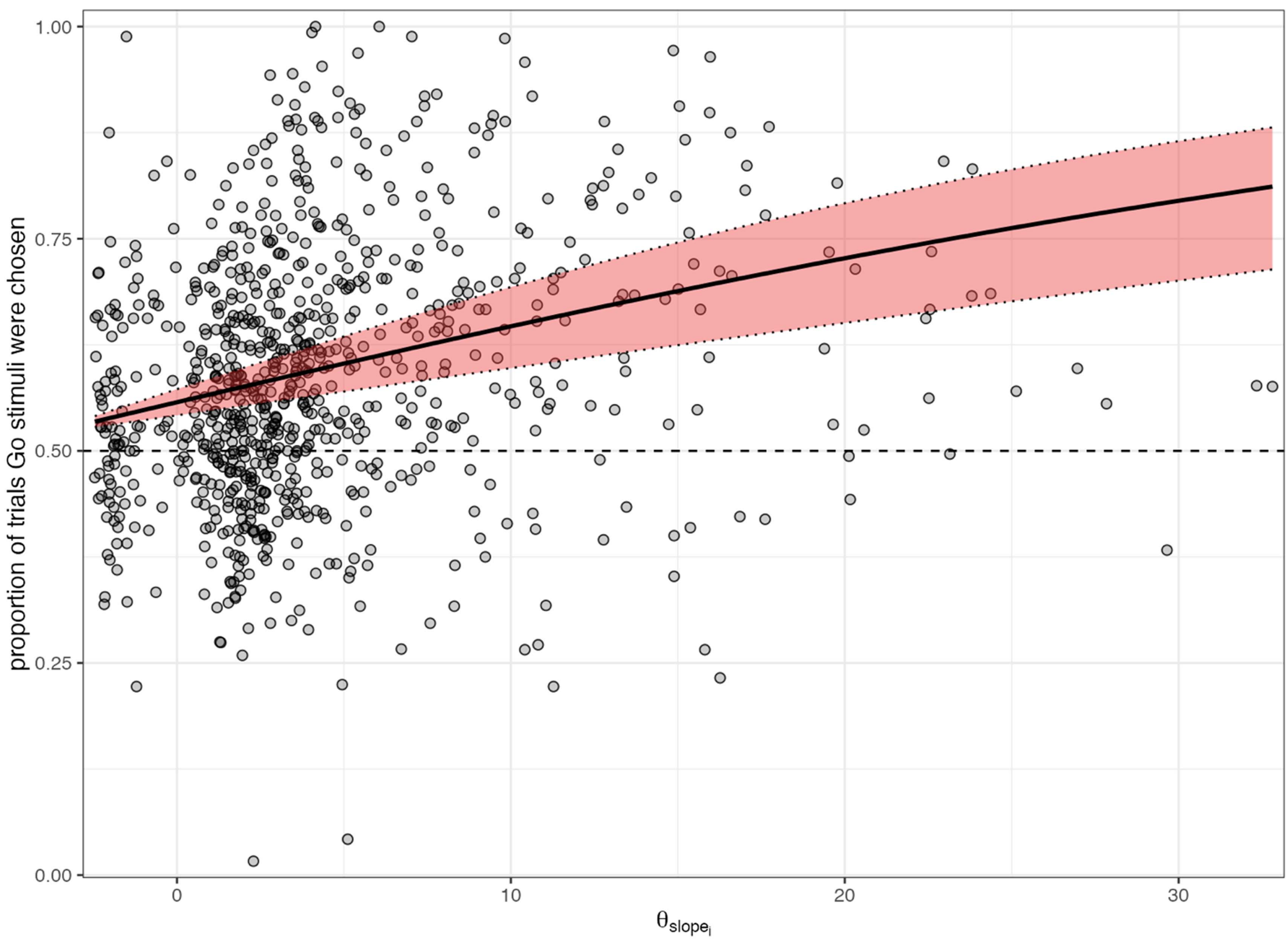
anticipatory responses, which we understood as a measure of the extent of learning about the stimulus achieved by the end point of CAT. To evaluate the association of the θ_{slope_i} parameter with preference change effect following CAT, we analyzed the proportion of probe trials in which participants chose the go over the no-go stimulus, as a per-participant linear function of θ_{slope_i} (mixed logistic regression model, including a random intercept and slope terms for the 28 experiments; see the Method section).

The meta-analysis showed θ_{slope_i} was positively associated with the preference modification effect—that is, participants with higher θ_{slope_i} parameter estimate, also demonstrated greater odds of choosing go stimuli ($OR = 1.04$, 95% CI [1.02, 1.05], $Z = 4.99$, $p < .001$; two-sided mixed model logistic regression; Figure 4). The model's intercept was significantly greater than zero, that is, even when extrapolated to a very low θ_{slope_i} value, the model forecasted enhanced preference for go stimuli (intercept odds = 1.26, 95% CI [1.18, 1.34], $Z = 7.21$, $p < .001$).

The estimated variances associated with the fixed-effect terms, random-effect terms, and residuals were used to evaluate two R^2 scores for the generalized (logistic) linear mixed mode: a marginal $R^2_{\mathrm{GLMM(m)}}$ (representing the relative proportion of variance associated with the fixed effects) and a conditional $R^2_{\mathrm{GLMM(c)}}$ score (representing the relative proportion of variance associated with the fixed effects and random effects; see more details in the Method section). The θ_{slope_i} fixed effect accounted for $R^2_{\mathrm{GLMM(m)}} = 0.31$ of the variance. With the random intercept and slopes, the fixed and random effects combined accounted for $R^2_{\mathrm{GLMM(c)}} = 0.83$ of the overall variance.

Additional Model Validation. In a post hoc analysis (see Supplemental Materials), we examined the explanatory power of θ_{slope_i} above and beyond a simpler RT-based marker. We used the proportion of anticipatory responses (RTs which were faster than the top 1% of RTs in the first run) each participant made as an alternative marker, and found that this marker was also associated

Figure 4
Meta-Analysis Results—Computational Marker and Preference Modification Effect



Note. Participants that transitioned faster to anticipatory responses during cue-approach training (larger θ_{slope_i} estimates) also demonstrated stronger preference modification effect (proportion of trials go stimuli were chosen). Trend line and surrounding red margins represent estimated preference modification effect and 95% CI, respectively (mixed model logistic regression). Dots represent individual participants. CI = confidence interval. See the online article for the color version of this figure.

with subsequent probe choices. Furthermore, θ_{slope_i} showed no significant contribution above and beyond this simpler RT-based marker.

In an additional post hoc analysis, we used an alternative Bayesian model in which RTs were modeled using log-normal distributions instead of Gaussians (see Supplemental Material). This new model better captured the right-tail shape of RTs (Supplemental Figure S13) and replicated similar correlation pattern between θ_{slope_i} parameter estimate and go stimuli choices during probe ($OR = 1.06$, 95% CI [1.04, 1.07], $Z = 4.92$, $p < .001$; see Supplemental Figure S14).

Stimulus-Specific Learning Parameter. While the previous model aggregated the data within participants, in fact, each participant has learned the go cue association with several different go stimuli. Thus, it is possible that learning was not uniform across the entire training stimuli, meaning every participant had a slightly different learning rate for each of the go stimuli she or he encountered. In an additional model, we aimed to expend our model by examining behavior on a more granular scale—focusing on the variability in RT and choice for individual go stimuli within participant. Accordingly, the computational model was elaborated by decomposing the per-subject learning parameter θ_{slope_i} into a set of subject- and stimulus-specific parameters $\theta_{slope_{i,s}}$, indicating the transition speed to anticipatory RTs of each go stimulus$_s$ encountered by participant$_i$ during training. The model was otherwise identical in design (see the Method section). This reanalysis of the data was conducted after the collection and analysis of the following two studies, which are reported in the current work.

The reanalyzed model converged around slightly different parameter estimates: Estimated early anticipatory RT distribution was characterized with lower mean ($\mu_1 = 89.45$, 95% CI [85.48, 93.42], $\sigma_{\varepsilon_1} = 276.77$ ms, 95% CI [274.36, 279.20]), late cue-dependent responses were similar ($\mu_2 = 287.62$ ms, 95% CI [287.15, 288.10], $\sigma_{\varepsilon_2} = 74.65$ ms, 95% CI [74.25, 75.04]), and the

$\theta_{slope_{i,s}}$ parameters tended to be lower but with greater variability (θ_{slope} = 2.17, 95% CI [2.00, 2.33], $\sigma_{\theta_{slope}}$ = 5.93, 95% CI [5.74, 6.13]; see Supplemental Figure S3).

Similar to the previous logistic regression analysis with per-participant parameter, using the current analysis' per-stimulus parameters also significantly predicted choices of the corresponding stimuli in the subsequent probe phase (intercept: odds = 1.44, 95% CI [1.36, 1.53], Z = 11.12, $p < .001$; $\theta_{slope_{i,s}}$ OR = 1.02, 95% CI [1.01, 1.04], Z = 3.56, $p < .001$; two-sided mixed-model logistic regression; see Supplemental Figure S4). Thus, go stimuli which were characterized with larger proportion of anticipatory responses were also chosen more during the probe phase. The models' fixed effects accounted for $R^2_{GLMM(m)} = 0.01$ of the total variance and the fixed effect with mixed effects accounted for $R^2_{GLMM(c)} = 0.78$ of the total variance.

However, this overall effect comprises both between-participant and within-subject, between-stimulus contributions. To examine the unique contribution of $\theta_{slope_{i,s}}$ above and beyond participant-level θ_{slope_i} parameter (which was evaluated in the previous analysis), in a third logistic regression model, we included both per-participant and per-stimulus parameters from the two different Stan models, as independent variable (and their random effects; see methods) to predict choices during probe. Our results indicate that each parameter provided some unique contribution above and beyond the other parameter (intercept: odds = 1.29, 95% CI [1.21, 1.38], $p < .001$; stimulus-level $\theta_{slope_{i,s}}$: OR = 1.04, 95% CI [1.02, 1.06], Z = 3.56, $p <$ .001; participant-level θ_{slope_i}: OR = 1.02, 95% CI [1.00, 1.03], Z = 2.71, p = .007). Thus, stronger learning of anticipatory RTs, both per-individual and per-stimulus within individual, was associated with increased odds of choosing go stimuli.

Interim Discussion

Examining RT patterns during CAT revealed for the first time a distinct time-dependent RT pattern, in which participants gradually transition from slow cue-dependent responses to rapid anticipatory RTs. Using Bayesian modeling, we were able to quantify this RT transition pattern with stable parameter estimates, of which the θ_{slope_i} parameter provided a promising computational marker to evaluated individualized differences in learning during training. Furthermore, as expected, this computational marker was also found to be associated with the preference modification effect in the subsequent binary choice probe phase, as participants with a more robust learning marker also demonstrated stronger preference modification. Examining the data at a stimulus-level learning parameter demonstrated an improved explanatory power for the variability in probe choices, above and beyond the wider participant-level parameter. Since the computational marker was calculated based on measurements that preceded the probe phase, it holds potential to serve as a predictive marker for future preference modification.

Several challenges were noted in modeling the original experimental data with the new computational model, largely because the original experimental procedures were not designed with the present analyses in mind. First, in the training task, go cue onsets were derived based on performance using a staircase procedure. Participants who performed well and responded rapidly were presented in the next trials with a more challenging cue, appearing later into the trial, while participants who performed poorly were exposed to earlier occurring (less difficult) cues. Consequently, anticipatory responses were not time-locked to the same event as the late cue-dependent responses. While a model that uses RTs time-locked to the stimulus onset could potentially model anticipatory responses well, using a model that is agnostic of the cue-onset time would result in difficulty to differentiate between cue-dependent and anticipatory responses due to the changing cue onset. For example, a response at 750 ms could be attributed to cue-dependent response of a participant with very poor performance which was presented with a cue at 500 ms, or an anticipatory response of a well-performing participant that encountered a challenging 800-ms cue-onset delay. Therefore, we analyzed RTs time-locked to the cue onset, which improved the distinction between cue-dependent and anticipatory responses, at the expense of less accurate tracking of the anticipatory response distribution shape.

Some lack of fit of the predicted data could also be indicative that the model's simplistic assumption that anticipatory RTs were normally distributed around the same mean for all participants was unlikely. This challenging effect may have led to some counterintuitive results; for instance, due to high variability in anticipatory responses, the fitted model predicted that very slow responses (RT effective > 500) would be more likely in the fastest learners' late runs, compared with earlier runs or of slower learners. In an additional post hoc analysis (performed after the initial write-up of this article), we modeled RT using a log-normal distribution, which better captured the skewed right-tail shape of RTs, which replicated the conclusions of the current design. Nonetheless, the large heterogeneous data structure (with over 800 participants from 28 different experiments) provided a computational challenge when trying to fit more complex models (e.g., with random effects per participant on this parameter or using nonnormal distributions). The current model was selected as a reasonable. The fact that using an additional alternative modeling approach resulted in similar conclusions as those with the preregistered Gaussian model provided important evidence that the conclusions of the present study were stable above and beyond implementation choices.

Furthermore, the task instructions of CAT specifically mentioned that participants are required to respond *after* cue onset. An additional crucial factor which could not be controlled in the present study was how rigorously participants complied with this instruction. It is possible that some participants might have learned the stimulus–cue association well and could have produced anticipatory responses but were reluctant to deviate from the task's guidelines. If so, there θ_{slope_i} would not reliably reflect their actual learning of the task. It is important to note that although the logistic regression results showed a significant positive association of choices with participant-level and even stimulus-level individualized learning parameters, significant preference change remains even when the model was extrapolated for a no learning effect. The effect sizes reported above are quite modest and in a post hoc analysis, they were not found significantly better compared to predicting choices with a simpler RT-based marker. This could suggest that there are additional factors affecting choices, which are not otherwise accounted for by the computational marker. The small effect sizes might be due to the inaccurate measurement of additional constructs within the computation parameter, such as the instruction adherence and interaction with performance-based dynamic cue onset, which were proposed here.

Despite these drawbacks, our unique model provided a prospective computational marker for learning and subsequent preference changes within the only variable behavior in the task: button pressing. Based on the promising results of Study 1 meta-analysis and considering the challenges raised above in fitting the computational model to the CAT task data, in Study 2, we devised a novel design of the CAT procedure, which was tested with two independent experiments. The first was a smaller preliminary study (n = 25) which was followed by a larger preregistered replication study (n = 59). In these new experiments, we addressed the limitations of the previous study design. We also modified the training design to answer a more directional hypothesis which does not only examine correlational association, but also aim to provide evidence for more consequential directionality, by inducing a differential behavioral change.

Study 2: Novel CAT Design—Preliminary and Replication Experiments

Following the results of Study 1, we devised a novel experimental design to test two preregistered experiments—a preliminary experiment and a larger replication experiment. Based on the conclusions from Study 1, the CAT procedure and instructions were altered to optimize the task for the Bayesian computational framework and manipulate behavior in accordance with the hypothesized cognitive mechanism.

In the training task, the go cue-onset time was fixed at 850 ms from stimulus onset, and participants were clearly instructed that they were permitted to make anticipatory responses before cue onset. Based on the conclusions from the meta-analysis study, we introduced within the training task a manipulation of the predictability of the stimulus–go cue-contingency, of the go stimuli, half of the stimuli were always associated with the go cue (100% contingency condition as in all previous CAT studies), while the rest of the go stimuli were only associated in half of the presentations with the go cue (50% contingency condition). The rest of the stimuli were never associated with a go cue (no-go stimuli; see the Method section for detailed description of the new design). We expected that this manipulation would affect the θ_{slope_i} parameter and subsequently also manipulate the preference modification effect.

We hypothesized that in the 50% contingency condition, learning the association of go stimuli with the go cue would be more challenging and would therefore be identified by a lower θ_{slope_i} learning parameter. We also hypothesized that manipulating learning efficacy during training would induce a differential preference modification effect in the subsequent probe phase, with more robust preference modification for go stimuli trained in the 100% contingency condition, compared with the more challenging 50% contingency condition.

The new task design was first tested in a preliminary (not preregistered) study with $n = 20$ valid participants, aimed to evaluate the efficacy of the new task and the replicability of the meta-analysis conclusions. Following the preliminary experiment, we ran an additional preregistered direct replication experiment with a larger sample size (n = 59), a size chosen based on the preliminary experiment results. As the two experiments were identical in design, they are reported together in all sections.

Method

Stimuli

In the two novel experiments, the stimuli set included images of 80 unfamiliar faces from the Siblings data set (Vieira et al., 2014), 40 man and 40 women characters. As in previous CAT studies with this stimulus-set (Salomon et al., 2018, 2019), the images were processed in Photoshop, so that all stimuli were cropped to identical size (400 × 500 pixels), with the character's pupils positioned at the same spatial coordinates and a homogeneous gray background. The characters poised similar neutral expression and had minimal salient artificial characteristics such as jewelry, makeup, or distinct facial hair. Previous CAT studies with these stimuli found that a similar preferences enhancement for both high- and low-value faces which were associated with a go cue during CAT (Salomon et al., 2018, 2019). Meaning, CAT had similar effect on preferences both for stimuli of initial low value and of initial high value. Using face stimuli allowed us to pool together a larger sample of stimuli of different initial values, under the reasonable assumption that the effect would show similar patterns across the different initial value categories.

During the CAT task, a visual go cue was associated with some of the face stimuli. The go cue was identical to the one used in a previous study (Salomon et al., 2018), it comprised a semitransparent Gabor image (38 × 38 pixels, $\alpha = .7$), which appeared at the center of the screen, on top of the face stimuli.

Procedure

Like previous CAT experiments, the two novel preliminary and replication experiments included three phases: a baseline preference evaluation task, a modified CAT task, and a post-training probe phase which examined preference modification.

Initial Preference Evaluation Task. To evaluate initial preferences, participants underwent a binary forced choice task between random pairs of stimuli, as in previous CAT experiments with nonconsumable stimuli (Salomon et al., 2018, 2019). The task included 400 unique choice trials (meaning, no choice between the same two faces repeated more than once), during which each of the 80 face stimuli was presented exactly 10 times. Each choice trial lasted 3,000 ms, of which participants were given a 2,000-ms time window to make their choice. Choice trials were followed by a 500-ms confirmation screen, showing a green frame around the chosen stimulus, and a fixation cross, which was presented for the remaining trial duration as interstimulus interval (ISI), for at least 500 ms. In case no choice was made during the 2,000-ms time-window, a screen saying "You must respond faster" appeared for 500 ms, followed by a 500-ms ISI.

Binary choices were transformed into individual preference scores using the Colley ranking algorithm (Colley, 2002). Stimuli were ranked based on each participant's individual initial preferences from the highest value (1) to lowest value (80). Ranks were used to categorize the stimuli into 10 equal-size value groups (each containing eight stimuli; ranks 1–8, ranks 9–16, etc.). The value categories and internal ranks within each category were used to allocate conditions in the subsequent training and probe task, ensuring initial values were balanced across 100% and 50% contingency conditions, as well as across go and no-go stimuli within each value category, which would be pitted against each other in the subsequent probe phase.

CAT. The CAT task consisted of 20 training runs, in each run, all 80 stimuli were presented in a random order for 1,000 ms each, with a 500-ms ISI. In 30% of trials a visual semitransparent go cue appeared 850 ms following the stimulus onset for 100 ms at the center screen position, on top of the face stimulus. Unlike previous CAT experiments, the go cue onset did not change throughout the training task. Thus, actual RT was consistent with effective RT (RT from cue onset).

Three go association conditions were included in this modified version of the CAT—16 stimuli were always presented with the go cue (go stimuli, 100% contingency condition), 16 stimuli were associated with the go cue during half of the presentations (go stimuli, 50% contingency condition), and the rest of the stimuli were never followed with a go cue (no-go stimuli). Participants were instructed to respond to go stimuli by pressing a keyboard button as fast as they could, before the face stimulus disappears. Participants were told in the instruction of the three go cue contingencies that they may respond, when they see the cue or even when they anticipate the cue will appear shortly (see the task instruction as presented to the participants in Supplemental Figure S5). Unlike previous CAT studies, participants were explicitly told of the go contingency and were thus encouraged to initiate anticipatory responses preceding the actual cue onset, while maintaining high accuracy by only responding when they have sufficient confidence that a cue will follow.

Stimuli were allocated with a go contingency condition based on initial preferences. Stimuli were categorized into 10 equal sized initial value categories, each containing eight stimuli. The stimuli in the highest and lowest initial value categories were allocated to be all no-go stimuli. Of the middle eight value categories, four were allocated to be 100% contingency and 50% contingency condition category. Within each of the eight middle categories, half of the stimuli were go stimuli, and the other half were no-go stimuli. The role allocation between categories and within them was designed, so that initial values were balanced across the groups for example, within the second value categories, stimuli ranked 9, 12, 13, and 16 were allocated to be go stimuli, while stimuli ranked 10, 11, 14, and 15 were allocated to be no-go stimuli (thus, each group had the same mean rank of 12.5; see illustration in Figure 5). The conditions allocation by initial value was counterbalanced across participants using four design combinations. In each combination, the value group allocation for 100% contingency was switched with that of 50% contingency stimuli, and within each value group, the allocation to go/no-go stimuli were flipped.

All go stimuli of the 50% contingency condition were associated with the go cue in the last two runs as well as in eight additional runs (in total—10 out of 20 runs). This was done for potential future implementation of the task in a magnetic resonance imaging scanner, which have not been done at the time of this article write-up.

Probe. In the probe phase, the preference modification effect was examined using a binary forced choice task, in which participants chose their preferred stimulus of pairs of go versus no-go stimuli, with similar initial value. Each go stimulus was pitted against the four no-go stimuli within the same initial value category (e.g., go stimuli ranked 9, 12, 13, and 16 were pitted against no-go ranked 10, 11, 14, and 15). Participants had 1,500 ms to make their choice, which was followed by a 500-ms confirmation feedback (or a message prompting faster response, in case no choice was made within the allocated 1,500 ms), and an ISI of varying duration (1,000–9,500 ms, $M = 3{,}000$ ms), drawn from a truncated exponential distribution with 100 ms precision. Like in previous CAT experiments with nonconsumable stimuli (and in contrast to CAT experiments with consumable stimuli), choices in the probe phase were not incentive compatible. Previous studies showed that the CAT effect was consistent both across incentive compatible choices of consumable stimuli as well as in nonincentive compatible choices (Salomon et al., 2018).

Choices of the 100% contingency condition and 50% contingency condition were merged across the different value categories and analyzed with a mixed logistic regression model. We hypothesized that participants would choose the go stimuli over no-go stimuli, and that this preference modification effect would be more robust for the 100% contingency condition. We also hypothesized that individual θ_{slope_i} parameters from computational modeling of the training

Figure 5
Go Stimuli Allocation Illustration

Value category	Ranks	Contingency for Go stimuli
1 (highest)	1 – 8	All NoGo
2	9 – 16	100%
3	17 – 24	50%
4	25 – 32	50%
5	33 – 40	100%
6	41 – 48	100%
7	49 – 56	50%
8	57 – 64	50%
9	65 – 72	100%
10 (lowest)	73 – 80	All NoGo

Value rank	Go allocation
9	100% Go
10	NoGo
11	NoGo
12	100% Go
13	100% Go
14	NoGo
15	NoGo
16	100% Go

Note. Both allocations to go contingency condition (100% vs. 50%) and go/no-go stimuli were balanced based on initial subjective value. Go stimuli of 50% contingency were associated in 10 of the 20 training runs with the go cue. Different designs were counterbalanced across participants by switching between 100% and 50% contingency categories positions and within category go/no-go positions.

task would predict which participants would demonstrate stronger preference modification effects.

Participants

In the preliminary experiment and replication experiment $n = 20$ and $n = 59$, valid participants completed the experiment and were included in the analysis, respectively. The sample size of the preliminary experiments was based on the minimal sample size used in previous CAT studies. The sample size required for the replication sample was based on a power analysis using the results of the preliminary experiment. A sample size of $n = 59$ was expected to be sufficient to achieve 95% power to detect a significant correlation effect ($\alpha = .05$) between probe choices and the θ_{slope_i} parameter estimates, in both 100% and 50% contingency conditions. The power analysis and the resulting sample sized was documented in the preregistration and its Supplemental Materials (https://osf.io/nwr4v).

Three quality assurance measurements were used as exclusion criteria, based on previous studies with CAT (Botvinik-Nezer et al., 2020; Salomon et al., 2018, 2019)—(1) low variability of Colley scores in the initial preference evaluation task (which indicate intransitive choice pattern), (2) proportion of false alarm during training, and (3) proportion of missed go trials during training. In each experiment, we excluded participants with extremely low transitivity, high false alarm, or high miss rate (defined as 3*SD* from the group mean). The exclusion criteria were preregistered for the replication experiment, along with the planned sample size.

In the preliminary experiment, no participant was excluded due to the mentioned above exclusion criteria (transitivity score: M = 0.205, 3*SD* cutoff = 0.122, min valid score = 0.137; false alarm rate: M = 3.62%, 3*SD* cutoff = 15.80%, max valid score = 15.27%; miss rate: M = 6.33%, 3*SD* cutoff = 24.13%, max valid score = 21.04%). In the replication experiment, five participants were excluded due to these preregistered exclusion criteria: Two participants had low transitivity score, one participant had high rate of false alarm, and two participants had high rate of missed go trials (transitivity score: M = 0.213, 3*SD* cutoff = 0.141, M_{valid} = 0.216, min valid score = 0.143; false alarm rate: M = 4.74%, 3*SD* cutoff = 39.00%, M_{valid} = 3.49%, max valid score = 37.59%; miss rate: M = 7.59%, 3*SD* cutoff = 40.22%, M_{valid} = 5.97%, max valid score = 38.75%).

Analysis

Like the procedural design, the analyses of the two experiments were also identical, and generally resembled that of the meta-analysis study, introducing some analysis improvements which could not be applied in the meta-analysis.

CAT Computational Model. The general goal and design of the Bayesian computational framework in the two new experiments resembled that of Study 1. Participants' RTs were modeled as a mixture of two Gaussian distributions—one distribution of late, cue-dependent responses, and another distribution of earlier anticipatory responses. A $\theta_{t,i}$ mixture proportion determined the probability of making an anticipatory response by a participant$_i$ at trial$_t$. The $\theta_{t,i}$ probability was modeled as a monotonic function of time, using as an individual rate parameter. Unlike the meta-analysis model, in Study 2, we introduced two distinct contingency conditions, thus for each participant two θ_{slope_i} parameters were modeled, namely, one for the 100% contingency condition and one for the 50% contingency condition. Another discrepancy between the meta-analysis model and the model of Study 2 was the introduction of an individual parameter for anticipatory responses center (μ_{1_i}), which allowed flexibility in modeling individual differences in anticipatory responses onsets (Equation 9). This model improvement was possible, thanks to the more homogeneous nature of the data. Each experiment (preliminary and replication) was modeled separately. See the complete model and priors in Supplemental Code S3.

$$\begin{aligned}
&\text{RT}_{t,i} = N(\mu_{1_i}, \sigma_{\varepsilon_1})(\theta_{t,i}) + N(\mu_2, \sigma_{\varepsilon_2})(1-\theta_{t,i})\\
&\theta_{t,i} = \Phi[-3.1 + I_{\text{cond.}_{t,i}}\theta_{100\%\text{slope}_i}\text{Run}_{t,i} + (1 - I_{\text{cond.}_{t,i}})\theta_{50\%\text{slope}_i}\text{Run}_{t,i}]\\
&\theta_{100\%\text{slope}_i} \sim N(\theta_{100\%}, \sigma_{\theta_{100\%}}); \theta_{50\%\text{slope}_i} \sim N(\theta_{50\%}, \sigma_{\theta_{50\%}});\\
&\quad \mu_{1_i} \sim N(\mu_1, \sigma_{\mu_1})\\
&\mu_{1_i} < \text{CueOnset} + 100, \mu_2 > \text{CueOnset};\\
&t\text{–trial index}; i\text{–participant index}\\
&I_{\text{cond.}_{t,i}} - \text{condition index (100\% versus 50\%)}.
\end{aligned} \tag{9}$$

In an additional exploratory analysis (not preregistered), we aimed to create an even more accurate learning computational marker by modeling a $\theta_{\text{slope}_{i,s}}$ for every stimulus$_s$ that participant$_i$ was trained with. Assuming that some variability in training speed could be measured not only between participants but also within each participant. Since the actual stimuli, which were allocated to be go stimuli varied between participants (i.e., for each participant, the condition and role of a certain image stimulus was different), we assumed no mutual information is shared between the same stimulus index of different participants. Stimuli and participants were treated as random effects (Equation 10).

$$\begin{aligned}
&\text{RT}_{t,i} = N(\mu_{1_i}, \sigma_1)(\theta_{t,i}) + N(\mu_2, \sigma_2)(1-\theta_{t,i})\\
&\theta_{t,i} = \Phi(-3.1 + I_{\text{stim.}_{t,i}}\theta_{\text{slope}_{i,s}}\text{Run}_{t,i})\\
&\mu_{1_i} < \text{CueOnset}, \mu_2 > \text{CueOnset};\\
&t\text{–trial index}; i\text{–participant index}; s - \text{stimulus index}\\
&I_{\text{stim.}_t} - \text{stimulus index (for each contingency condition separately)}.
\end{aligned} \tag{10}$$

In the process of testing this novel approach, the model did not converge when presented with the full data which included both 100% and 50% contingency conditions. However, when trained twice, using each contingency condition as an independent data set, the model did converge. See the complete model and priors in Supplemental Code S4.

Probe. As in previous CAT studies, preference modification was evaluated following CAT. Participants' preferences for go over no-go stimuli were categorized into two conditions corresponding with the CAT conditions—of 100% contingency and 50% contingency conditions. Trials of the different value groups were pooled together and analyzed with a mixed-model logistic regression. In a post hoc analysis, we validated that the initial value had no significant impact on the conclusions of the analysis (see Supplemental Materials).

Based on the results of the meta-analysis, we hypothesized that better learning (manifested as a larger θ_{slope_i} parameter estimate) would induce a stronger preference modification effect. Thus, we also hypothesized that a stronger preference modification effect would be observed for the 100% contingency condition, compared with the 50% contingency condition. To examine this hypothesis, we ran a one-sided mixed-model logistic regression with additional fixed and participant-based random explanatory variables of the contingency condition (Equation 11).

$$\begin{aligned}\text{Choice} \sim 1 &+ \text{Contingency} \\ &+ (1 + \text{Contingency}|\text{Participant}). \end{aligned} \tag{11}$$

Probe Choice Prediction Based on Individual Learning Parameter. Most importantly, we hypothesized that using the θ_{slope_i} parameter estimate for learning would predict future preference modification effect observed during the subsequent probe phase. To test this hypothesis, we introduced the θ_{slope_i} as an additional independent variable in the mixed-model logistic regression. Choices were modeled using contingency condition (both as fixed effect and as a random slope between participants), θ_{slope_i}, and an interaction term. This model is equivalent to modeling choices using an intercept and θ_{slope_i} slope separately for each contingency condition with a random intercept modeled within participants (Equation 12).

In an additional (not preregistered) exploratory analysis, we modeled a stimulus-specific $\theta_{\text{slope}_{i,s}}$ parameter. We aimed to examine if modeling a learning parameter at a more detailed stimulus level (extending the preregistered participant-level θ_{slope_i} parameter) would serve as a more accurate predictor of future preference modification. For this aim, we merged the $\theta_{\text{slope}_{i,s}}$, which were estimated for each contingency with the probe data. Choices were modeled with two additional models using a $\theta_{\text{slope}_{i,s}}$ as an independent variable. In the first model, $\theta_{\text{slope}_{i,s}}$ parameter estimates were used instead of θ_{slope_i} estimates. Since $\theta_{\text{slope}_{i,s}}$ varied within participant, a random slope effect for $\theta_{\text{slope}_{i,s}}$ was also added (Equation 13). In another model, probe choices were explained using a full model with both $\theta_{\text{slope}_{i,s}}$ and θ_{slope_i} (Equation 14). This model was used to examine whether $\theta_{\text{slope}_{i,s}}$ provided additional explanatory power in a likelihood ratio test for nested models (comparing the model of Equation 14 with the nested model of Equation 12).

$$\text{Choice} \sim 1 + I_{\text{cond.}} \times \theta_{\text{slope}_i} + (1 + I_{\text{cond.}}|\text{Participant}), \tag{12}$$

$$\begin{aligned}\text{Choice} \sim 1 &+ I_{\text{cond.}} \times \theta_{\text{slope}_{i,s}} \\ &+ (1 + I_{\text{cond.}} \times \theta_{\text{slope}_{i,s}}|\text{Participant}),\end{aligned} \tag{13}$$

$$\begin{aligned}&\text{Choice} \sim 1 + I_{\text{cond.}} \times \theta_{\text{slope}_i} + I_{\text{cond.}} \times \theta_{\text{slope}_{i,s}} \\ &\quad + (1 + I_{\text{cond.}} \times \theta_{\text{slope}_{i,s}}|\text{Participant}) \\ &i\text{–participant index}; s - \text{stimulus index} \\ &I_{\text{cond.}} - \text{contingency condition indicator IV.}\end{aligned} \tag{14}$$

When results with $\theta_{\text{slope}_{i,s}}$ were analyzed, we found that some participants were fitted a $\theta_{\text{slope}_{i,s}}$ estimates of extremely small variability ($SD < 0.1$, meaning their $\theta_{\text{slope}_{i,s}}$ estimates were effectively fixed; see Supplemental Figure S6). In such cases a mixed model with $\theta_{\text{slope}_{i,s}}$ random slope (Equations 13 and 14), resulted in convergence warnings (this phenomenon was not observed in Study 1). Thus, for these analyses, we excluded all participants which demonstrated such low variability of $\theta_{\text{slope}_{i,s}}$ estimates in either one of the two contingency conditions. This resulted in exclusion of nine participants in the preliminary experiment, and 14 participants in the replication experiment (out of 20 and 59 participants, respectively). For the nested model likelihood ratio analysis (comparing a full model to model without $\theta_{\text{slope}_{i,s}}$ independent variables), both models were examined using the same subset of 11 and 45 valid participants.

Effect Size Estimation for Mixed Models With R^2_{GLMM}. As in the meta-analysis study, we report for the logistic regression mixed models two R^2 estimates, $R^2_{\text{GLMM(m)}}$ and $R^2_{\text{GLMM(c)}}$, denoting respectively the marginal and conditional R^2 effect of the generalized linear model (Bartoń, 2020; Johnson, 2014; Nakagawa et al., 2017; Nakagawa & Schielzeth, 2013).

Results

Reaction Times in the Training Task

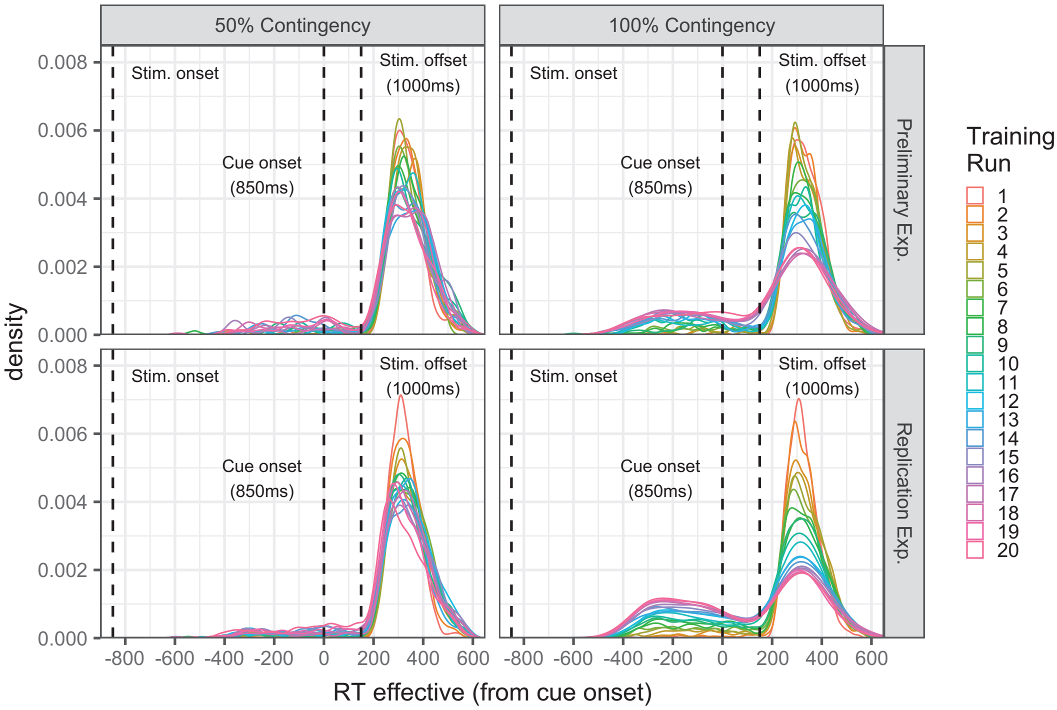
As in the meta-analysis study, RT patterns of the training task were examined. Before examining RT patterns, we removed go trials, where participants did not respond at all (preliminary exp. = 6.33%. replication exp. 5.97% of go trials) or trials with an unlikely fast RT (RT < 250 ms, preliminary exp. = 0.31%, replication experiment 0.21% of trials). In total, 636 (of 9,600; 6.62%) and 1,747 (of 28,320; 6.17%) of trials were excluded from the preliminary and replication experiments, respectively. Since the proportion of missed trials was more significant in Study 2, we validated in a post hoc analysis that eliminating these trials did not have a significant impact on the research conclusions (see Supplemental Material).

In both experiments, we were able to replicate the RT patterns observed in the meta-analysis study—while in early training runs, participants mostly relied on cue-dependent RTs, as training progressed, an increased portion of RTs were of earlier anticipatory RTs (Figure 6). Examining mean (*SD*) effective RT at Run 1, revealed similar RTs in all contingency conditions—preliminary exp.: $M_{50\%\text{Cont.}} = 326.88$ ms (85.13 ms), $M_{100\%\text{Cont.}} = 336.60$ ms (64.39 ms); replication exp.: $M_{50\%\text{Cont.}} = 322.49$ ms (63.68 ms), $M_{100\%\text{Cont.}} = 319.90$ ms (78.88 ms). As training progressed, participants were able to anticipate the cue and respond before cue onset. Examining the differential pattern between the 100% contingency condition and the 50% contingency condition showed, as expected, that participants initiated more anticipatory responses in the 100% contingency condition, compared with the 50% contingency condition, resulting in faster mean RTs at Run 20—preliminary exp.: $M_{50\%\text{Cont.}} = 253.27$ ms (187.65 ms), $M_{100\%\text{Cont.}} = 175.55$ ms (245.5 ms); replication exp.: $M_{50\%\text{Cont.}} = 256.57$ ms (188.94 ms), $M_{100\%\text{Cont.}} = 81.39$ ms (271.99 ms); see Figure 6.

Computational Marker of Learning. As in the meta-analysis study, RTs within each condition were modeled with a time-dependent mixture model. In both the preliminary and replication samples, the computational model converged around similar hyperparameter estimates (see MCMC trace plot and posterior simulations fit in Supplemental Figures S6 and S7, respectively). Cue-dependent responses were characterized with later mean RT and smaller variance, compared with earlier anticipatory responses (preliminary exp.: $\mu_1 = 735.07$, 95% CI [677.14, 792.87], $\sigma_{\varepsilon_1} = 140.14$, 95% CI [130.13, 153.49], $\mu_2 = 1{,}190.23$, 95% CI [1,188.46, 1,192.03], $\sigma_{\varepsilon_2} = 78.11$, 95% CI [76.78, 79.47];

Figure 6

Density Plots of Effective RT Distribution (Time-Locked to Cue Onset) in Study 2



Note. Preliminary experiment on top row and replication experiment on bottom row: The go cue onset was fixed at 850 ms from stimulus onset—vertical dashed lines represent the trial onset (time 0), cue onset, and trial offset (850 ms and 1,000 ms from trial onset, respectively). As training progressed (indicated by the training run in color), responses were less homogeneous following cue onset, with increased proportion of anticipatory responses. The transition to anticipatory responses was more robust in the 100% contingency condition (right column) compared with the 50% contingency condition (left column). RT = reaction time; Exp. = experiment.

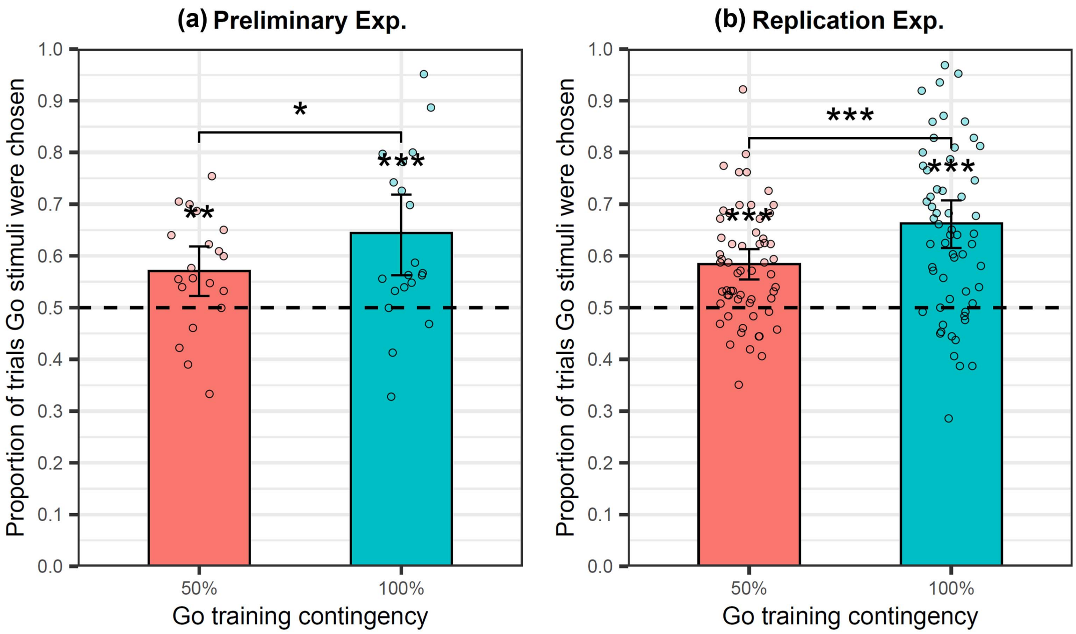
replication exp.: μ_1 = 740.44, 95% CI [710.10, 771.76], σ_{ε_1} = 140.35, 95% CI [135.79, 144.74], μ_2 = 1,187.46, 95% CI [1,186.25, 1,188.68], σ_{ε_2} = 80.30, 95% CI [79.46, 81.15]). As the cue onset was fixed in Study 2 at 850 ms, the mean cue-dependent RTs of approximately 1,190 ms corresponded with an effective RTs center of approximately 340 ms in both experiments. Anticipatory RT mean varied considerably between participants, as identified by the σ_{μ_1} parameter, which modeled between-participant variability in μ_{1_i} parameter estimates (preliminary exp.: σ_{μ_1} = 110.81, 95% CI [73.29, 166.69]; replication exp.: σ_{μ_1} = 110.99, 95% CI [90.92, 136.34]).

Transition from cue-dependent to anticipatory responses, as captured by θ_{slope_i} parameter estimates of the preliminary experiment, indicated faster transition in the 100% contingency condition, compared with the 50% contingency condition ($\theta_{\text{slope100\%}}$ = 1.95, 95% CI [1.03, 2.75], $\theta_{\text{slope50\%}}$ = 0.90, 95% CI [−0.01, 1.68]; mean difference in θ_{slopes} = 1.04, 95% CI [-0.17, 2.23]), with prominent variability between participants ($\sigma_{\theta_{\text{slope100\%}}}$ = 1.98, 95% CI [1.45, 2.67], $\sigma_{\theta_{\text{slope50\%}}}$ = 1.79, 95% CI [1.21, 2.57]). This effect was replicated even more distinctly in the larger replication experiment ($\sigma_{\theta_{\text{slope100\%}}}$ = 3.25, 95% CI [2.54, 3.89], $\sigma_{\theta_{\text{slope50\%}}}$ = 1.54, 95% CI [1.11, 1.92]; mean difference in θ_{slopes} = 1.70, 95% CI [0.90, 2.47]; $\sigma_{\theta_{\text{slope100\%}}}$ = 2.82, 95% CI [2.37, 3.42], $\sigma_{\theta_{\text{slope50\%}}}$ = 1.45, 95% CI [1.16, 1.82]). The 95% credible interval of the difference between the two conditions θ_{slope_i} parameters indicated a general trend in the preliminary experiment which was distinct in the replication sample.

Preference Modification in Probe Task. Based on the results of the meta-analysis in Study 1, we hypothesized that manipulating the go cue-contingency would correspondingly manipulate the preference modification effect, manifested as more pronounced preference modification for stimuli in the 100% contingency condition, compared with the 50% contingency condition.

As expected, in the preliminary experiment, participants chose go stimuli over no-go stimuli above chance level, both in the 50% contingency condition (prop. = 57.11%, Z = 2.85, p = .002, odds = 1.33, 95% CI [1.09, 1.62]; one-sided logistic mixed model), as well as in the 100% contingency condition (prop. = 64.45%, Z = 3.40, p < .001, odds = 1.81, 95% CI [1.29, 2.56]). The preference modification effect was more robust in the 100% contingency condition, than in the 50% contingency condition (Z = 1.92, p = .028, OR = 1.36, 95% CI [0.99, 1.87]). These effects were replicated in the larger replication sample (50% contingency condition: prop. = 58.41%, Z = 5.50, p < .001, odds = 1.40, 95% CI [1.24, 1.59]; 100% contingency condition: prop. = 66.31%, Z = 6.41, p < .001, odds = 1.97, 95% CI [1.60, 2.42]; condition difference effect: Z = 3.81, p < .001, odds = 1.40, 95% CI [1.18, 1.67]; one-sided logistic mixed model), see Figure 7.

Figure 7
Probe Results



Note. Proportion of trials participants chose go stimuli over no-go stimuli of similar initial value, in the preliminary experiment (a) and replication experiment (b). Dots represent individual participants, error bars represent 95% CI based on a mixed-model logistic regression. Participants demonstrated enhance preference both for go stimuli in the 50% contingency condition and to a larger extent in the 100% condition. Dashed line represents 50% chance level. CI = confidence interval; Exp. = experiment. See the online article for the color version of this figure.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$; one-sided mixed-model logistic regression.

In post hoc analyses, we examined the impact of initial subjective value on the reported effects by controlling for the initial value difference within each probe choice as well as by examining the interaction of the contingency effect with value categories. In all the analyses, our effects of interest remained consisted, that is, we observed overall enhanced preferences for go over no-go stimuli, and a more robust preference for go stimuli within the 100% contingency condition (see Supplemental Materials).

Prediction of Choice Behavior in the Probe Using the Computational Model. Most importantly, we aimed to examine whether the preference modification effect could be predicted using the individually fitted θ_{slope_i} parameter estimated for each participant. Adding the individual learning parameter as an additional independent variable to the logistic regression model revealed that indeed the θ_{slope_i} parameter estimated from the training task could be used to predict preference modification in the subsequent probe task. In both experiments, a significant contribution was found for θ_{slope_i} in the 50% contingency condition (preliminary exp.: log-*OR* = 0.12, $Z = 2.30$, $p = .011$, *OR* = 1.13, 95% CI [1.02, 1.26]; replication exp.: log-*OR* = 0.10, $Z = 2.39$, $p = .008$, *OR* = 1.11, 95% CI [1.02, 1.20]; one-sided mixed-model logistic regression) as well as in the 100% contingency condition (preliminary exp.: log-*OR* = 0.26, $Z = 3.59$, $p < .001$, *OR* = 1.29, 95% CI [1.12, 1.49]; replication exp.: log-*OR* = 0.11, $Z = 3.49$, $p < .001$, *OR* = 1.12, 95% CI [1.05, 1.19]). No significant differences were found between the slopes under the two conditions (preliminary exp.: log-*OR* = 0.13, $Z = 1.52$, $p = .13$, *OR* = 1.14, 95% CI [0.96, 1.36]; replication exp.: log-*OR* = 0.01, $Z = 0.26$, $p = .79$, *OR* = 1.01, 95% CI [0.93, 1.10]; two-sided mixed-model logistic regression), see Figure 8.

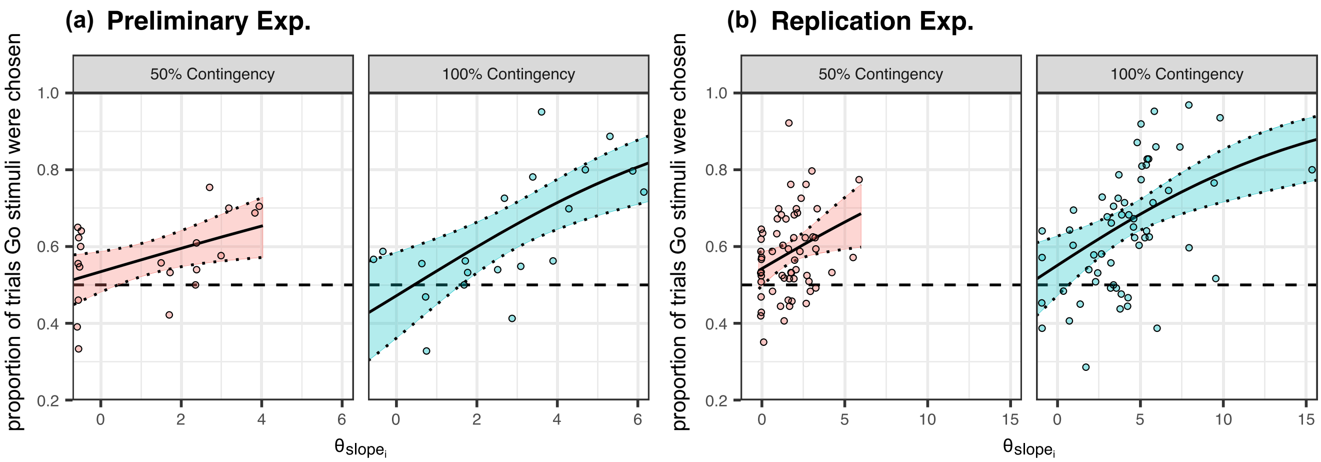
Examining the sizes of estimated fixed-effects, random-effects, and residual variances revealed that the model accounted for a large portion of the variance in participants' choice pattern in both experiments, both when examining the contribution of the fixed effects only (preliminary exp.: $R^2_{\text{GLMM(m)}} = 0.421$; replication exp.: $R^2_{\text{GLMM(m)}} = 0.223$), and furthermore when examining the joint contribution of fixed and random effects (preliminary exp.: $R^2_{\text{GLMM(c)}} = 0.869$; replication exp.: $R^2_{\text{GLMM(c)}} = 0.872$).

To examine the overall effect of the change in contingency condition (50% vs. 100%) above and beyond θ_{slope_i}, we used a likelihood ratio test to compare the full logistic regression model with a restricted (nested) model that did not contain contingency fixed effects (contingency intercept and slope interaction; restriction of two parameters). In both experiments, comparing the two models revealed no significant contribution of contingency above and beyond θ_{slope_i} (preliminary exp.: $\Delta\text{AIC}_{(\text{rest. - full})} = -1.64$, $\chi^2_{(2)} = 2.36$, $p = .307$; replication exp: $\Delta\text{AIC}_{(\text{rest. - full})} = -3.61$, $\chi^2_{(2)} = 0.39$, $p = .824$; two-sided likelihood ratio test).

Post Hoc Model Validation

Like in Study 1, in a post hoc analysis, we fitted a non-Gaussian model (log-normal distributions) and found a similar fit to actual RT

Figure 8
Association of Probe Choices With the Training Computational Marker in Study 2



Note. Larger θ_{slope_i} parameter estimates were positively associated with subsequent preference modification in probe phase, both in the preliminary experiment (a) and in the larger replication experiment (b). The 100% contingency condition was characterized with larger θ_{slope_i} values and respectively also stronger preference modification effect. Trend lines and surrounding color margins represent estimated preference modification effects and 95% CI, respectively (mixed-model logistic regression). Dots represent individual participants. Horizontal dashed line represents 50% chance level. CI = confidence interval; Exp. = experiment. See the online article for the color version of this figure.

data (see Supplemental Figure 16), as well as positive association of the θ_{slope_i} parameter estimates with subsequent probe choice both in preliminary experiment (50% condition: log-*OR* = 0.03, $Z = 2.45$, $p = .014$, $OR = 1.03$, 95% CI [1.02, 1.04]; 100% condition: log-*OR* = 0.04, $Z = 3.75$, $p < .001$, $OR = 1.04$, 95% CI [1.03, 1.51]; two-sided mixed model logistic regression) and in the replication study in 50% condition only (50% condition: log-*OR* = 0.014, $Z = 2.12$, $p = .038$, $OR = 1.01$, 95% CI [1.008, 1.02]; 100% condition: log-*OR* = 0.003, $Z = 0.71$, $p = .48$, $OR = 1.003$, 95% CI [0.998, 1.008]; two-sided mixed-model logistic regression; see full report in Supplemental Material). We also performed a post hoc analysis in which we compared the explanatory power of θ_{slope_i} with a simpler marker using the proportion of anticipatory responses each participant made (i.e., the proportion of RTs which were faster than the top 1% of RTs in the first run; see Supplemental Material). In contrast to the previous meta-analysis results, in the two Study 2 experiments, we found that θ_{slope_i} had a significant predictive power, above and beyond a simpler marker which is based on the proportion of anticipatory responses, participants made during training (see full report in Supplemental Material).

Examining the Interplay With Choice RT. One hypothesis raised suggested that CAT might impact preference for go stimuli by encouraging faster impulsive choices (Veling, Chen, et al., 2017; Veling, Lawrence, et al., 2017). To examine the hypothesis that θ_{slope_i} impacts choices via faster automated response to go stimuli, we introduced two new exploratory analyses which examined whether θ_{slope_i} correlated with faster RT during probe. We found no evidence supporting that θ_{slope_i} was associated with choice RT during probe, preliminary experiment—50% contingency choices: $b = 4.56$, $t(18) = 0.18$, $p = .859$, 100% contingency: $b = 9.27$, $t(18) = 0.42$, $p = .682$; replication experiment—50% contingency choices: $b = 1.18$, $t(57) = 0.09$, $p = .928$, 100% contingency: $b = 3.79$, $t(57) = 0.57$, $p = .568$; two-sided linear mixed model.

In an exploratory analysis, we further examined whether the predictive power of θ_{slope_i} remained consistent above and beyond the time it took participants to make their probe choices. For that aim, we introduced an additional regressor of the choice RT (in seconds) to the logistic regression analysis and found that choice RT often had a significant explanatory power. Nonetheless, θ_{slope_i} was consistently found to be predictive of choosing go stimuli, above and beyond choice RT (see detailed results in Supplemental Material).

Stimulus-Specific Learning Parameter. In an additional analysis (mentioned but not detailed in the preregistration), we extended this analysis to stimulus-level effects. Each go stimulus$_s$ presented to participant$_i$ was fitted a $\theta_{slope_{i,s}}$ parameter, which was expected to capture within-participant variability in learning—that is, model for difference in learning of stimuli within a contingency condition. We aimed to use these participant and stimulus-specific learning parameters as independent variables in a mixed-model logistic regression (with a random intercept, random slope for contingency, and random slope for $\theta_{slope_{i,s}}$ independent variables, see the Method section for full description of the statistical models).

Fitting a full Stan model for both stimuli of 50% and 100% contingency simultaneously did not converge. We hypothesized that the full model, with Two Conditions × 16 Stimuli × *n* Participants Interdependent $\theta_{slope_{i,s}}$ parameters was too complex to resolve using an MCMC process. Thus, we decided to reduce complexity by eliminating the participant-level effects on the mean RT for the early anticipatory responses (estimating only group-level μ_1, without participant-level μ_1 parameters as was done in the previous meta-analysis study). We further split the data based on contingency condition and analyzed each contingency as an independent data set. Using this approach resolved the convergence issues. This solution for an unexpected issue was not preregistered or planned prior to data analysis.

Examining the association between the $\theta_{slope_{i,s}}$ parameter estimates and the participant-level θ_{slope_i} parameter estimates of the previous model revealed high similarity (preliminary exp.: $r = 0.84$; replication exp.: $r = 0.80$). In addition, participants with negative θ_{slope_i} parameter estimates had very small variance ($SD < 0.1$) in their $\theta_{slope_{i,s}}$ estimates (see Supplemental Figure S8). When such $\theta_{slope_{i,s}}$ were introduced in the mixed-model logistic regression, the low within-participant variability caused convergence warnings. Therefore, all participants with $\theta_{slope_i} < 0.2$ (which is equivalent to $SD < 0.1$ of $\theta_{slope_{i,s}}$) in either contingency condition, were excluded from the analysis. This resulted in the exclusion of nine participants in the preliminary experiment and 14 participants in the replication experiment.

Analyzing the remaining 11 and 45 participants in the preliminary and replication experiment resulted in similar conclusions, as found with the participant-level θ_{slope_i} parameter estimate. Overall, in both contingency conditions, a significant positive association was found between $\theta_{slope_{i,s}}$ and preference of go stimuli over no-go stimuli (preliminary experiment—50% contingency: $Z = 4.34$, $p < .001$, $OR = 1.33$, 95% CI [1.17, 1.52]; 100% contingency: $Z = 2.85$, $p = .002$, $OR = 1.21$, 95% CI [1.06, 1.38], overall model $R^2_{GLMM(m)} = 0.269$ $R^2_{GLMM(c)} = 0.631$; replication experiment—50% contingency: $Z = 3.14$, $p < .001$, $OR = 1.09$, 95% CI [1.03, 1.15]; 100% contingency: $Z = 5.09$, $p < .001$, $OR = 1.09$, 95% CI [1.05, 1.12], overall model $R^2_{GLMM(m)} = 0.147$, $R^2_{GLMM(c)} = 0.475$; one-sided mixed-model logistic regression; Figure 9). No significant differences were found between the slopes of the two conditions (preliminary exp.: $Z = -1.35$, $p = .18$, $OR = 0.91$, 95% CI [0.79, 1.04]; replication exp.: $Z = -0.08$, $p = .94$, $OR = 1.00$, 95% CI [0.95, 1.05]; two-sided mixed-model logistic regression).

The contribution of stimulus-level $\theta_{slope_{i,s}}$ parameter to predicting preferences was examined by comparing a full model, containing both stimulus-level $\theta_{slope_{i,s}}$ and participant-level θ_{slope_i} individualized learning parameters to a restricted model containing only θ_{slope_i} as independent variable. The addition of $\theta_{slope_{i,s}}$ had a significant contribution both in the preliminary experiment ($\Delta AIC = 59.65$, $\chi^2_{(9)} = 77.65$, $p < .001$; likelihood ratio test) as well as in the replication experiment ($\Delta AIC = 93.97$, $\chi^2_{(9)} = 111.97$, $p < .001$).

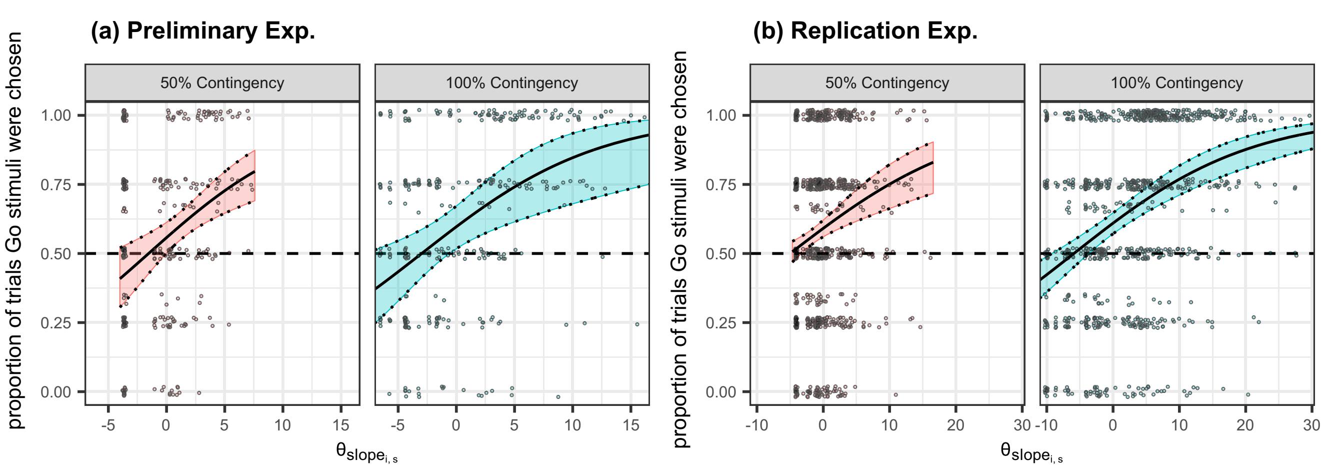
Interim Discussion

Based on the conclusions of the meta-analysis study, a novel preference modification procedure was designed and tested in Study 2. In the new design, two cue-contingency conditions (50% and 100% contingency) were tested within-participant, which we hypothesize would manipulate the difficulty of learning the stimulus–cue association during the training phase of the CAT paradigm. The introduction of the two contingency conditions allowed us to experimentally impact the θ_{slope_i} individualized computational marker of learning, which in turn provided a reliable predictor for the subsequent preference modification effect in the probe phase. As expected, we found a more robust preference modification effect for go stimuli in the 100% contingency condition, compared with the 50% contingency condition. The fact that a directed experimental intervention in the preceding training phase induced a differential effect in the probe phase provides support for a directional causal impact of training on preference change. Considering these results, it is plausible to conclude that the association demonstrated in Study 1 between the individualized learning parameter and behavioral change might represent a causal relationship, where improved learning in the training phase induces stronger behavioral change in the subsequent probe phase.

Furthermore, between-participant variability in the θ_{slope_i} computational marker was also predictive of between-participant difference in preference modification, as measured during the subsequent probe phase. Participants with a more robust θ_{slope_i} also demonstrated stronger preference for go stimuli over no-go stimuli of similar initial value. This effect was observed both in the 100%

Figure 9

Association of Probe Choices With Stimulus-Level Computational Markers in Study 2



Note. Larger $\theta_{slope_{i,s}}$ estimates parameter estimates were positively associated with subsequent preference modification in probe phase, both in the preliminary experiment (a) and in the larger replication experiment (b). Dots represent individual stimuli (small vertical jitter was added to better visualize high-density areas). Trend lines and surrounding color margins represent estimated preference modification effect and 95% CI, respectively (mixed-model logistic regression). Horizontal dashed line represents 50% chance level. CI = confidence interval; Exp. = experiment. See the online article for the color version of this figure.

and 50% contingency conditions, with no significant difference in the effect between the two conditions. This suggests that θ_{slope_i} had a similar predictive pattern for both conditions, one unit increase in θ_{slope_i} resulted in a similar increase in the odds of choosing go over no-go, above and beyond the contingency condition. However, as the 100% contingency condition was characterized with larger θ_{slope_i} parameter estimates, so was the probe phase characterized with greater odds of choosing go stimuli (i.e., stronger preference modification effect). Taken together, these results provide evidence for a novel mean to both predict individual differences in learning and preference modification between participants as well as means to manipulate this preference modification effect. All the nonexploratory results in the preliminary experiment were carefully detailed in a preregistration and were all replicated in the larger replication experiment, which included the exact same procedure and analysis pipelines. It is important to note that using a larger sample size not only replicated the significance of the effects, but also the descriptive trends, resulting in outcomes of similar effect size.

In contrast to the post hoc analysis in Study 1, which found that the computational marker performed similarly well as a simple RT-based marker, in a post hoc analysis of the two new experiments, we did find that θ_{slope_i} captured a unique predictive signal above and beyond a simple proportion of anticipatory responses. These results reinforce our hypothesis, that by optimizing the training procedure for our computational model (by fixing the go cue-onset time and explicitly encouraging participant to make anticipatory responses), we were able to better capture learning signals using θ_{slope_i} marker.

It is interesting to note that descriptively the new design induced a strong preference modification effect in the 100% go contingency (group estimate of approximately 65% in both experiments), compared with other CAT and go/no-go experiments, where a more modest effect of 55%–60% preference bias for go stimuli was commonly observed (Aridan et al., 2019; Bakkour et al., 2016, 2017; Botvinik-Nezer et al., 2021; Salomon et al., 2018, 2019; Schonberg et al., 2014; Veling, Chen, et al., 2017; Zoltak et al., 2018). Without carefully controlling for all factors, it is hard to conclude whether this difference is significant, and which one of the changes in the experimental design induced this enhanced effect, however, it would be interesting to try and address this directly in future work.

In an exploratory analysis, we examined variation across stimuli as well as participants, focusing both on the variability in learning patterns between different participants, as well as within-participant variability in learning for different stimuli. Due to the technical requirements, such an in-depth analysis was applicable only for a smaller subset of our data; however, the results we found suggest that the prediction model is applicable in all levels of granularity, both at the individual participant level as well as in the individual stimulus level.

General Discussion

In the current work, we aimed to study and characterize the cognitive mechanisms underlying nonreinforced preference modification with CAT (Schonberg et al., 2014). While behavioral change is manifested in a binary probe phase, actual preference change occurs during the training phase where only individual items are presented. Based on previous studies that alluded to the potential involvement of internal nonreinforced learning in CAT, the current work aimed to use computational modeling to measure this internal process at the training phase with individual items and establish its causal role on the subsequent preference modification effect. The CAT task is a multiphase procedure, which includes an initial preference evaluation task, a nonreinforced training phase with a simple process of pressing a button and cue when pictures of individual items are presented, and a binary probe phase. Therefore, the simple training phase provided a unique opportunity to aggregate data from multiple studies to build a large corpus of training data.

Our work studies a current topic of wide interest on nonreinforced preference change (e.g., Chew et al., 2021) and moreover offers a unique and robust methodological approach. By using a simple task that could not be deciphered at the individual sample level, we aggregated 28 samples to perform computational modeling to describe the underlying processes. Then to complete the scientific process we preregistered, designed, and tested a manipulation to test our model.

In the first part of the current work, we examined previously collected data of $n = 840$ participants from 28 different CAT experiments in a meta-analysis and identified a distinct time-dependent RT pattern during the training task. We found that while in early stages of training, participants depended on the go cue to initialize responses (homogeneous RTs following cue onset), as training progressed, participants relied less on the go cue and started to generate fast anticipatory responses. This raised the hypothesis that a marker of transition from cue-dependent responses to early anticipatory responses could be indicative of internal nonreinforced learning and subsequent preference changes.

Using a Bayesian computational framework implemented with Stan, a statistical model of RT patterns was formulated. RTs were postulated to derive from a mixture model of two Gaussian distributions and a time-dependent mixture proportion, which accounted for the transition speed from late cue-dependent responses to early anticipatory responses. The model parameters were estimated with an MCMC algorithm, with a key participant-level θ_{slope_i} parameter, which modeled the individualized difference between participants in transition to anticipatory responses rate. This Bayesian parameter was used as a passive computational marker of learning. Examining the association between the individualized learning parameter and preference modification effect in the subsequent probe task, showed a positive association in which participants with more robust learning parameter estimates during training also demonstrated stronger preference modification effect in the subsequent probe phase (i.e., enhanced preference for the go stimuli over no-go stimuli of similar initial value). Thus, the θ_{slope_i} parameter, which was optimized to model RT patterns during training, was found to be good predictor of preference modification effect in a future probe task. An additional analysis modeling RT using a more finely tuned learning $\theta_{slope_{i,s}}$ parameter, fitted for the individual stimuli, further expended these findings, and demonstrated positive predictive power of $\theta_{slope_{i,s}}$ in evaluating preference modification effect on a stimulus-level basis, above and beyond the more general participant-level parameter.

Based on the results of Study 1, we derived a theoretical hypothesis with implications for shaping learning efficacy and consequently also the effect of training on preference behavior. If faster transition to anticipatory responses was related to learning, then hindering the ease of cue anticipation could also obstruct learning and resulting in less pronounces preference modification effect. To examine this hypothesis, we design a novel CAT procedure, which was tested in the second part of this work (Study 2), first with a preliminary experiment and then an additional larger preregistered replication experiment. Cue anticipation was manipulated by introducing two training conditions within the CAT task—in the 100%-contingency condition, the go stimuli perfectly anticipated go cue onset, while in the 50% contingency condition, the go stimuli only anticipated the go cue in half of the trials. As preregistered and predicted, manipulating cue-contingency during CAT induced the expected effects both in the training and in the probe task. During training, participant relied less on early anticipatory responses and transitioned slower to anticipatory responses in the 50% contingency condition. This effect was captured by the θ_{slope_i} estimates from the meta-analysis, which were larger in the 100% contingency condition compared with the 50% contingency condition. More importantly, in the subsequent probe phase, a consistent pattern of stronger preference modification effect was observed for the 100% contingency go stimuli. Thus, manipulating training difficulty affected future preference modification measured in the probe phase. When individual differences in preference modification were examined, the θ_{slope_i} marker accounted for a large portion of variability in choices, while no significant differences were found between the two contingency conditions above and beyond the θ_{slope_i} effect. This suggests that the more robust preference modification effect observed for the 100% contingency go stimuli was captured by the stark differences in θ_{slope_i} learning marker. Furthermore, the fact that intentionally manipulating the training procedure in Study 2 induced differential learning parameter estimates as well as differential behavioral change effect further supports the hypothesis of a causal nature of the association with the behavioral change. Thus, it is plausible to deduce that shaping training at the individual item level, which was captured by the individualized learning parameter, drove the induced differential behavior modification effect, as evaluated in the subsequent probe phase. It is important to note that although the cue-contingency was different between the two conditions, the actual exposure time to the stimuli was identical. Thus, the enhanced preference for go stimuli in the 100% contingency condition could not be accounted for by mere exposure effect (Zajonc, 1968).

What Is the Mechanism?

Our findings, showing a consistent link between motor planning during training at the individual items level and preference modification in the subsequent binary choice probe phase, correspond with previous works. These works showed that a rapid response is a crucial feature for preference modification with CAT (Bakkour et al., 2016), and neural finding showing that increased striatal and premotor activity during training, were associated with more robust preference modification effects in the probe phase (Salomon et al., 2019). Thus, the results suggest that learning efficacy in the training phase is manifested as stimulus-specific motor planning. Importantly, all participants were alert and attentive throughout the training phase, as demonstrated by the negligible rates of nonresponse. However, it is evident that mere attention is not sufficient to induce strong preference modification, but rather an attention that can be translated to action, suggesting a unique valuation pathway in the absence of external reinforcements, putatively based on parieto-frontal circuits involving attention with motor planning (Schonberg & Katz, 2020).

To produce an anticipatory response for a go stimulus, participants must also rely on memory. Thus, it is possible that the anticipatory response pattern captured in our data identified individualized differences in memory which in turn mediated a change in preferences (Weber & Johnson, 2009). This hypothesis is in line with previous findings with nonreinforced training, which identified a positive association between preferences modification and enhanced memory both at the participant level (Salomon et al., 2018) and at the stimulus level (Botvinik-Nezer et al., 2021; Chen et al., 2021). Since memory for the trained go stimuli was not examined in the current work, it would be interesting to examine in future experiments whether the θ_{slope_i} computational marker correlates with declarative memory and does it provide independent predictive power of subsequent nonreinforced preference modification above and beyond memory measurements. Future imaging studies could also examine the new task design with functional imaging to identify whether differential neural activation patterns characterize the two contingency conditions.

One question that is often raised in nonreinforced preference modification research is whether the behavioral effect of choosing go stimuli reflects an internal change in the value of the chosen go stimuli, or rather habitual automated responses to choose the go stimuli originating from mechanisms like operant conditioning. On the one hand, CAT experiments showed that not all cue–response associations result in enhanced preference for go stimuli—CAT was found to be less effective in enhancing preferences for stimuli of negative affective valence (Salomon et al., 2018) and low-value snack-food items (Schonberg et al., 2014), suggesting that the stimulus pertaining value interacts with the effect of CAT on preferences. In addition, nonchallenging CAT designs such as CAT where the cue starts with the stimulus onset or when the response association was made as a block of stimuli to which participants were required to respond also failed to induce preference change (Bakkour et al., 2016), which provide evidence that the mechanism impacting preferences requires more than a simple motor response association. On the other hand, some findings allude to the involvement of rapid impulsive decision-making mechanisms—CAT experiments in which participants were asked to make slow well-considered decisions found that the CAT effect on preferences diminished when participants made slow nonimpulsive choices. This result further resonates the finding that during the CAT probe phase, participants were faster in their choices when go stimuli were chosen (Veling, Chen, et al., 2017; Veling, Lawrence, et al., 2017).

To test these two competing notions that the CAT nonreinforced preference modification is driven by impulsive motor response habitual mechanisms versus internal value representation change, we examined in our data whether choice RT could account for the enhanced likelihood of choosing go over no-go stimuli during probe. Our results showed most of the times that indeed faster choice RTs were associated with increased likelihood of choosing go stimuli, in agreement with past findings. However, and more importantly, we found that the computational marker was not correlated with choice RT, and maintained its predictive power of choice above and beyond choice RT. Thus, our results indicate that while some aspects of nonreinforced preference change could be attributed to rapid impulsive-like choices, our proposed computational marker tracks an independent preference modification mechanism, which could putatively be indicative of value representation change. Future work could attempt to find behavioral or neural correlates which could corroborate this hypothesis.

Internal Reinforcement?

An explanation offered by the model for the modification of preferences revolves around the presence of an internal reinforcement process during the training phase. Following the initial exposure to stimulus–cue contingencies, participants attempted to respond based on the stimuli rather than the cue itself, displaying anticipatory responses. When participants correctly pressed earlier, an internal feedback mechanism was activated, further enhancing the learning process. This internal reinforcement can be seen as a broader mechanism for facilitating learning. It is challenging to differentiate this internal reinforcer from memory or attention processes, and our study does not aim to do so. Instead, the current model integrates these mechanisms and incorporates the internal reinforcer as a component of the learning process. Drawing on previous research and considering the correlation between brain regions associated with the reward system (Salomon et al., 2019), we propose that a reinforcement mechanism plays a role in preference change. Specifically, since no external feedback was provided during the training task, our hypothesis suggests that the reinforcement mechanism is likely to be internal in nature. The current work's unique contribution establishing a new passive marker for nonreinforced learning could be tested in future neuroimaging studies. These studies could shade light on whether the marker is associated more with parietal attentional mechanisms, temporal memory-related regions, or striatal reward and motor learning neural mechanisms.

In addition to a basic understanding of the cognitive mechanisms of nonreinforced preference change, identifying a learning marker based on early training runs could be used to design powerful predictive tools for learning efficacy, with relevance to other paradigms and situations that share the basic training/transfer structure of the current studies. A naive approach to evaluate learning efficacy during training tasks such as CAT might suggest introducing direct measurements of behavior change throughout the training process, for example, by asking participants to make active value-based choices (Hands, 2012; Vlaev et al., 2011). Relatedly, it has been shown that the mere act of choice could induce a longitudinal effect of enhanced preference for the chosen stimuli (Nakamura & Kawabata, 2013; Sharot et al., 2010, 2012) and introduced bias in the effect of preference modification following nonreinforced training (Chen et al., 2021). Therefore, probing preference within preference modification training is likely to alter the learning process and undermine its validity. Moreover, in more standard learning procedures, such as conditioning-based interventions, to examine the efficacy of learning, researchers observe the subject's response to the conditioned stimulus when the associated unconditioned stimulus is omitted (Barak et al., 2013; De Houwer et al., 2001; Frank et al., 2004, 2007; Joel, 2006; Kawa et al., 2016; LaBar et al., 1998; Tzschentke, 1998). Over repetition ofprobe tests in all these procedures might initiate a new learning procedure—where the conditioned stimulus is no longer associated with the unconditioned stimulus, which would eventually lead to extinction learning (M. E. Bouton, 1993; M. E. M. E. Bouton, 2004; Calcagnetti & Schechter, 1993; Myers & Davis, 2002). Thus, evaluating learning by active probing of the behavior change effect could interfere with the learning process. Importantly, a passive marker for learning that can predict behavioral change efficacy based on earlier training data does not evoke the drawbacks of introducing an active choice task to reveal preferences. Using a passive learning marker, which is evaluated based on independent training data that preceded the probe phase, overcomes this obstacle. The temporal primacy of θ_{slope_i} estimation could provide a prediction tool for future preference modification in CAT, without tainting the results with a direct evaluation of preferences using a choice task. An interesting hypothesis could anticipate that the 50% cue-contingency would resemble partial reinforcement learning schedule, and thus will have potentially higher long-term sustainability and resistance to extinction (Nevin, 1988). While the current work did not examine long-term maintenance of CAT effect and its association cue-contingency, this could be an interesting subject to test in future studies.

Real-Time Marker for Learning

The benefits of real-time individualized markers for learning could also be of great value for additional learning and behavior change procedures beyond CAT. In many learning tasks (including CAT), the behavioral change effect on choices is measured in a separate probe phase, following the training phase. For example, in experiments testing pain perception, participants are trained to associate neutral stimuli with differential level of painful heat stimulation before probing the impact of interventions such as drug versus placebo administration (Atlas et al., 2010; Koban et al., 2018). Studies examining habit formation may use lengthy free-operant learning protocols, in which participants repeatedly perform an action (such as pressing a button) to gain food rewards, in order to test in a later probe phase whether the participants demonstrate habitual behavior and continue to perform the action even when the reward is devalued (e.g., pressing to get food when satiated; Dolan & Dayan, 2013; Pool et al., 2021; E. Tricomi et al., 2009). Even outside the field of value-based decision-making, in the clinical psychological well-being domain, attention bias modification procedures use computerized attention training interventions similar in nature to CAT to treat depression and anxiety (Bar-Haim et al., 2007; Browning et al., 2012; Hakamata et al., 2010). In such experimental settings, where training is either unpleasant or exhausting, with plausible odds that learning would not be well-established if undertrained, monitoring learning efficacy individually in real time based on the training data could assist in optimizing training efficacy and efficiency. While the current work focuses on CAT, we assert that the basic mechanisms identified by the current work could be adapted to accommodate a need in a wider range of learning procedures. With an appropriate adaptation to the desired experimental design, this unique feature of a passive computational marker which is predictive of subsequent change, opens a new path for potential interventions that will allow more efficient training via monitoring and real-time feedback (Pardo, 2018; Tempelaar et al., 2015).

The current work's approach in modeling learning using RT could also be applicable in experimental designs which do not separate the probe from the training task, such as reinforcement learning tasks, where learning can be evaluated as it progresses on a trial-by-trial basis (Leong et al., 2017; Ratcliff, 1978; Schultz et al., 1997; Smith & Ratcliff, 2004). Since RT patterns could be indicative of the individual's confidence in her choice during reinforcement learning tasks such as probabilistic selection (Frank et al., 2004, 2007) or multiarm bandit task (Niv et al., 2012, 2015), incorporating RT data could improve reinforcement learning models and capture learning more accurately (Ballard & McClure, 2019; Fontanesi et al., 2019).

Limitations and Alternative Hypotheses

Some limitations were not directly tested in the current work and should be addressed in future studies. Primarily, in Study 2 of the current work, we demonstrated that learning and preference modification could be hindered using a more difficult/partial association procedure. However, our model also predicts that easing the association procedure could enhance learning efficacy and preference modification effect. Further work should empirically test this hypothesis, for example, by examining the effect of enhancing the go cue saliency or attention during predefined go stimuli (Armel et al., 2008; Krajbich et al., 2010; Shimojo et al., 2003) or by manipulating the temporal features of training schedule (Cepeda et al., 2008). Furthermore, by instructing participants to respond when they have enough confidence that a cue will follow, we might have introduced an undesired confound to the task in the subjectivity of confidence each participant needed to respond; that is, some participants might have learned well the association between a stimulus and the go cue but were reluctant to respond before they validate that in the 50% contingency condition a cue will appear. Although such confound does not undermine the validity of the findings, future studies should aim avoid it and examine a differential effect in which the difference between conditions reflects only task difficulty. One potential design could be by asking to respond to stimuli which at any point were associated with a cue or by clarifying in the instructions that a response based on a guess is also valid.

Another limitation of the current work could raise a concern for a confound which might undermine the ability to deduce causal relationship between learning and choice in our new experimental design. In the two experiments of Study 2, we found that participants demonstrated an enhanced preference modification effect for stimuli that were more consistently associated with the go cue. By maintaining a similar presentation time of both 50% and 100% contingency stimuli, we eliminated the confound of the standard mere-exposure effect (Zajonc, 1968). Participants viewed the stimuli in both conditions for the same time duration and putatively had to maintain high alertness to both type of stimuli, which both required a response with a high likelihood. However, one might argue that participants were more engaged with the 100% contingency stimuli, to which they pressed twice as much. Thus, an alternative explanation could claim that increased engagement (i.e., more press responses) is the true causal factor which induced a stronger preference modification effect, rather than the learning procedure identified by the computational marker.

Our statistical analysis provides evidence suggesting that the learning marker had a stronger impact on choices than the level of exposure. When analyzing the factors predicting choice patterns, the training condition had no explanatory power above and beyond the computational marker of learning. Nonetheless, a dedicated experimental design could aim to directly discern between these two competing hypotheses. One such proposed training design could present the 50% contingency stimuli twice as often as 100% contingency stimuli. This would result in a training design with more consistent association for the 100% contingency condition (similarly to the current design), while maintaining equal engagement levels between the two conditions (manifested as identical number of press responses in both training conditions) and an increased exposure to the 50% contingency stimuli, which would be presented for twice the duration of time as the 100% contingency condition stimuli. Our theory hypothesizes that despite the increased exposure and similar engagement with the 50% contingency stimuli in this design, participants will still demonstrate faster learning in the consistent 100% contingency condition and would thus show stronger preference modification effect for the 100% contingency condition stimuli. Future work could try to run this dedicated design and provide empirical evidence that would settle whether our proposed model of learning overcomes the mere exposure effect combined with more equal engagement level, which we did not fully address in the current work.

Although we primarily describe the effect of CAT on preferences as an "enhancement" of go stimuli, it is important to emphasize that this change is relative to the no-go stimuli in a binary choice setting. An alternative explanation proposes that CAT does not necessarily increase the intrinsic value of go stimuli but rather decreases the value of no-go stimuli. The latter concept aligns with findings from a go/nogo training study where high-calorie no-go food items elicited reduced neural activity in reward-related brain regions (Yang et al., 2023). However, our previous imaging study with CAT (Salomon et al., 2019) observed a positive correlation between enhanced striatal reward signals to go stimuli during training and a subsequent preference for those stimuli, without a corresponding decrease in neural response to no-go stimuli. This suggests that CAT may primarily increase the value of go stimuli without directly impacting the value of no-go stimuli. To further investigate this, future CAT experiments could directly compare preferences for no-go stimuli against untrained stimuli of similar initial value. The mere-exposure effect (Zajonc, 1968) predicts a preference change for familiar no-go stimuli, while the alternative hypothesis suggests a preference for the novel, untrained stimuli. Observing the latter would provide compelling evidence that the no-go association in CAT negatively impacts preferences.

Finally, we carefully documented the experimental choices in a preregistration before data were collected for an independent replication experiment in Study 2. Utilizing a Bayesian computational framework, such as the one used here, provided great flexibility required to fit a complicated theoretical model. We aimed to use a relatively straightforward model, with as similar features as possible in all studies. Future studies could take the liberty to use the openly accessible data from this work and try to improve the model by changing our current assumptions including distributions' shape, free parameters, and priors.

Constraints on Generality

This work is based on a meta-analysis of 28 studies performed in different countries on multiple samples. However, it is still limited to mainly human participant students from different universities (not all). This is especially the case for the target populations in Study 2 collected for this study at Tel Aviv University.

Conclusion

In conclusion, the current work laid the foundations for understanding and quantifying nonexternally reinforced learning at the individual item level of a simple training task via a novel Bayesian modeling approach of RT patterns during training. Using a large meta-analysis data set of previous CAT studies and two new original samples, we demonstrated a unique method to evaluate a learning marker at the individual item level with a robust predictive power of subsequent preference change. The task is simple with only non-reinforced button presses toward items that were presented multiple times prior to a subsequent probe phase. Thus, we propose that motor response patterns could provide a passive marker for value change, which does not require direct measurement of preferences.

References

Aridan, N., Pelletier, G., Fellows, L. K., & Schonberg, T. (2019). Is ventromedial prefrontal cortex critical for behavior change without external reinforcement? *Neuropsychologia*, *124*, 208–215. https://doi.org/10.1016/j.neuropsychologia.2018.12.008

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, *3*(5), 396–403. https://doi.org/10.1017/S1930297500000413

Atlas, L. Y., Bolger, N., Lindquist, M. A., & Wager, T. D. (2010). Brain mediators of predictive cue effects on perceived pain. *The Journal of Neuroscience*, *30*(39), 12964–12977. https://doi.org/10.1523/JNEUROSCI.0057-10.2010

Bakkour, A., Leuker, C., Hover, A. M., Giles, N., Poldrack, R. A., & Schonberg, T. (2016). Mechanisms of choice behavior shift using cue-approach training. *Frontiers in Psychology*, *7*, Article 421. https://doi.org/10.3389/fpsyg.2016.00421

Bakkour, A., Lewis-Peacock, J. A., Poldrack, R. A., & Schonberg, T. (2017). Neural mechanisms of cue-approach training. *NeuroImage*, *151*, 92–104. https://doi.org/10.1016/j.neuroimage.2016.09.059

Ballard, I. C., & McClure, S. M. (2019). Joint modeling of reaction times and choice improves parameter identifiability in reinforcement learning models. *Journal of Neuroscience Methods*, *317*, 37–44. https://doi.org/10.1016/j.jneumeth.2019.01.006

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Barak, S., Liu, F., Ben Hamida, S., Yowell, Q. V., Neasta, J., Kharazia, V., Janak, P. H., & Ron, D. (2013). Disruption of alcohol-related memories by mTORC1 inhibition prevents relapse. *Nature Neuroscience*, *16*(8), 1111–1117. https://doi.org/10.1038/nn.3439

Bartoń, K. (2020). *MuMIn: Multi-model inference* (R package Version 1.43.17). https://cran.r-project.org/web/packages/MuMIn/index.html

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226–232. https://doi.org/10.1002/bs.3830090304

Botvinik-Nezer, R., Bakkour, A., Salomon, T., Shohamy, D., & Schonberg, T. (2021). Memory for individual items is related to nonreinforced preference change. *Learning & Memory*, *28*(10), 348–360. https://doi.org/10.1101/lm.053411.121

Botvinik-Nezer, R., Salomon, T., & Schonberg, T. (2020). Enhanced bottom-up and reduced top-down neural mechanisms drive long-lasting non-reinforced behavioral change. *Cerebral Cortex*, *30*(3), 858–874. https://doi.org/10.1093/cercor/bhz132

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*(1), 80–99. https://doi.org/10.1037/0033-2909.114.1.80

Bouton, M. E. M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, *11*(5), 485–494. https://doi.org/10.1101/lm.78804

Browning, M., Holmes, E. A., Charles, M., Cowen, P. J., & Harmer, C. J. (2012). Using attentional bias modification as a cognitive vaccine against depression. *Biological Psychiatry*, *72*(7), 572–579. https://doi.org/10.1016/j.biopsych.2012.04.014

Calcagnetti, D. J., & Schechter, M. D. (1993). Extinction of cocaine-induced place approach in rats: A validation of the "biased" conditioning procedure. *Brain Research Bulletin*, *30*(5–6), 695–700. https://doi.org/10.1016/0361-9230(93)90102-H

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science*, *19*(11), 1095–1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

Chen, Z., Holland, R. W., Quandt, J., Dijksterhuis, A., & Veling, H. (2021). How preference change induced by mere action versus inaction persists over time. *Judgment and Decision Making*, *16*(1), 201–237. https://doi.org/10.1017/S1930297500008366

Chew, B., Blain, B., Dolan, R. J., & Rutledge, R. B. (2021). A neurocomputational model for intrinsic reward. *The Journal of Neuroscience*, *41*(43), 8963–8971. https://doi.org/10.1523/JNEUROSCI.0858-20.2021

Clithero, J. A., & Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience*, *9*(9), 1289–1302. https://doi.org/10.1093/scan/nst106

Colley, W. (2002). *Colley's bias free college football ranking method: The Colley matrix explained*. Princeton University.

Collins, A. G. E., & Frank, M. J. (2014). Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological Review*, *121*(3), 337–366. https://doi.org/10.1037/a0037015

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*(6), 853–869. https://10.1037/0033-2909.127.6.853

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*, *26*(4), 1099–1121. https://doi.org/10.3758/s13423-018-1554-2

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16311–16316. https://doi.org/10.1073/pnas.0706111104

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science*, *306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., Ernst, M., & 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

Hands, D. W. (2012). Foundations of contemporary revealed preference theory. *Erkenntnis*, *78*(5), 1081–1108. https://doi.org/10.2139/ssrn.2068724

Joel, D. (2006). The signal attenuation rat model of obsessive–compulsive disorder: A review. *Psychopharmacology*, *186*(4), 487–503. https://doi.org/10.1007/s00213-006-0387-2

Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R^2_{GLMM} to random slopes models. *Methods in Ecology and Evolution*, *5*(9), 944–946. https://doi.org/10.1111/2041-210X.12225

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: Consensus and controversy. *Neuron*, *63*(6), 733–745. https://doi.org/10.1016/j.neuron.2009.09.003

Kawa, A. B., Bentzley, B. S., & Robinson, T. E. (2016). Less is more: Prolonged intermittent access cocaine self-administration produces incentive-sensitization and addiction-like behavior. *Psychopharmacology*, *233*(19–20), 3587–3602. https://doi.org/10.1007/s00213-016-4393-8

Koban, L., Kusko, D., & Wager, T. D. (2018). Generalization of learned pain modulation depends on explicit learning. *Acta Psychologica*, *184*, 75–84. https://doi.org/10.1016/j.actpsy.2017.09.009

Kraemer, P. M., Weilbächer, R. A., Mechera-Ostrovsky, T., & Gluth, S. (2022). Cognitive and neural principles of a memory bias on preferential choices. *Current Research in Neurobiology*, *3*, Article 100029. https://doi.org/10.1016/j.crneur.2022.100029

Krajbich, I. (2019). Accounting for attention in sequential sampling models of decision making. *Current Opinion in Psychology*, *29*, 6–11. https://doi.org/10.1016/j.copsyc.2018.10.008

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. https://doi.org/10.1038/nn.2635

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13852–13857. https://doi.org/10.1073/pnas.1101328108

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. *Neuron*, *20*(5), 937–945. https://doi.org/10.1016/S0896-6273(00)80475-4

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-8.

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*, *93*(2), 451–463. https://doi.org/10.1016/j.neuron.2016.12.040

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *Karolinska directed emotional faces* [Database record]. APA PsycTests. https://doi.org/10.1037/t27732-000

Myers, K. M., & Davis, M. (2002). Behavioral and neural analysis of extinction. *Neuron*, *36*(4), 567–584. https://doi.org/10.1016/S0896-6273(02)01064-4

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society, Interface*, *14*(134), Article 20170213. https://doi.org/10.1098/rsif.2017.0213

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nakamura, K., & Kawabata, H. (2013). I choose, therefore I like: Preference for faces induced by arbitrary choice. *PLOS ONE*, *8*(8), Article e72071. https://doi.org/10.1371/journal.pone.0072071

Nevin, J. A. (1988). Behavioral Momentum and the Partial Reinforcement Effect. *Psychological Bulletin*, *103*(1), 44–56. https://doi.org/10.1037/0033-2909.103.1.44

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. https://doi.org/10.1016/j.jmp.2008.12.005

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *The Journal of Neuroscience*, *35*(21), 8145–8157. https://doi.org/10.1523/JNEUROSCI.2978-14.2015

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *The Journal of Neuroscience*, *32*(2), 551–562. https://doi.org/10.1523/JNEUROSCI.5498-10.2012

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452–454. https://doi.org/10.1126/science.1094285

Pardo, A. (2018). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education*, *43*(3), 428–438. https://doi.org/10.1080/02602938.2017.1356905

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051

Pool, E. R., Gera, R., Fransen, A., Perez, O. D., Cremer, A., Aleksic, M., Tanwisuth, S., Quail, S., Ceceli, A. O., Manfredi, D. A., Nave, G., Tricomi, E., Balleine, B., Schonberg, T., Schwabe, L., & O'Doherty, J. P. (2021). Determining the effects of training duration on the behavioral expression of habitual control in humans: A multilaboratory investigation. *Learning & Memory*, *29*(1), 16–28. https://doi.org/10.1101/lm.053413.121

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556. https://doi.org/10.1038/nrn2357

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. https://doi.org/10.1126/science.1100301

Rindfleisch, A., & Inman, J. J. (1998). Explaining the familiarity-liking relationship: Mere exposure, information availability, or social desirability? *Marketing Letters*, *9*(1), 5–19. https://doi.org/10.1023/A:1007958302123

Salomon, T., Botvinik-Nezer, R., Gutentag, T., Gera, R., Iwanir, R., Tamir, M., & Schonberg, T. (2018). The cue-approach task as a general mechanism for long term non-reinforced behavioral change. *Scientific Reports*, *8*, Article 3614. https://doi.org/10.1038/s41598-018-21774-3

Salomon, T., Botvinik-Nezer, R., Oren, S., & Schonberg, T. (2019). Enhanced striatal and prefrontal activity is associated with individual differences in nonreinforced preference change for faces. *Human Brain Mapping*, *41*(4), 1043–1060. https://doi.org/10.1002/hbm.24859

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, *5*(17), 61–71. https://doi.org/10.2307/2548836

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

Schonberg, T., & Katz, L. N. (2020). A neural pathway for nonreinforced preference change. *Trends in Cognitive Sciences*, *24*(7), 504–514. https://doi.org/10.1016/j.tics.2020.04.002

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Sharot, T., Fleming, S. M., Yu, X., Koster, R., & Dolan, R. J. (2012). Is choice-induced preference change long lasting? *Psychological Science*, *23*(10), 1123–1129. https://doi.org/10.1177/0956797612438733

Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological Science*, *21*(9), 1231–1235. https://doi.org/10.1177/0956797610379235

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322. https://doi.org/10.1038/nn1150

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168. https://doi.org/10.1016/j.tins.2004.01.006

Staddon, J. E. R., & Cerutti, D. T. (2003). Operant conditioning. *Annual Review of Psychology*, *54*, 115–144. https://doi.org/10.1146/annurev.psych.54.101601.145124

Stan Development Team. (2020). *RStan: The R interface to Stan* (R package Version 2.19.3). https://mc-stan.org/

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Tanji, J. (1996). New concepts of the supplementary motor area. *Current Opinion in Neurobiology*, *6*(6), 782–787. https://doi.org/10.1016/S0959-4388(96)80028-6

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior*, *47*, 157–167. https://doi.org/10.1016/j.chb.2014.05.038

Tom, G., Nelson, C., Srzentic, T., & King, R. (2007). Mere exposure and the endowment effect on consumer decision making. *The Journal of Psychology*, *141*(2), 117–125. https://doi.org/10.3200/JRLP.141.2.117-126

Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *The European Journal of Neuroscience*, *29*(11), 2225–2232. https://doi.org/10.1111/j.1460-9568.2009.06796.x

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron*, *41*(2), 281–292. https://doi.org/10.1016/S0896-6273(03)00848-1

Tzschentke, T. M. (1998). Measuring reward with the conditioned place preference paradigm: A comprehensive review of drug effects, recent progress and new issues. *Progress in Neurobiology*, *56*(6), 613–672. https://doi.org/10.1016/S0301-0082(98)00060-4

Veling, H., Chen, Z., Tombrock, M. C., Verpaalen, I. A. M., Schmitz, L. I., Dijksterhuis, A., & Holland, R. W. (2017). Training impulsive choices for healthy and sustainable food. *Journal of Experimental Psychology: Applied*, *23*(2), 204–215. https://doi.org/10.1037/xap0000112

Veling, H., Lawrence, N. S., Chen, Z., van Koningsbruggen, G. M., & Holland, R. W. (2017). What is trained during food Go/No-Go training? A review focusing on mechanisms and a research agenda. *Current Addiction Reports*, *4*(1), 35–41. https://doi.org/10.1007/s40429-017-0131-5

Vieira, T. F., Bottino, A., Laurentini, A., & De Simone, M. (2014). Detecting siblings in image pairs. *The Visual Computer*, *30*(12), 1333–1345. https://doi.org/10.1007/s00371-013-0884-3

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. A. (2011). Does the brain calculate value? *Trends in Cognitive Sciences*, *15*(11), 546–554. https://doi.org/10.1016/j.tics.2011.09.008

Voigt, K., Murawski, C., & Bode, S. (2017). Endogenous formation of preferences: Choices systematically change willingness-to-pay for goods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(12), 1872–1882. https://doi.org/10.1037/xlm0000415

Weber, E. U., & Johnson, E. J. (2009). Constructing preferences from memory. In S. Lichtenstein, & P. Slovic (Eds.), *The construction of preference* (pp. 397–410). Cambridge University Press. https://doi.org/10.1017/CBO9780511618031.022

Weilbächer, R. A., Krajbich, I., Rieskamp, J., & Gluth, S. (2021). The influence of visual attention on memory-based preferential choice. *Cognition*, *215*, Article 104804. https://doi.org/10.1016/j.cognition.2021.104804

Yang, Y., Morys, F., Wu, Q., Li, J., & Chen, H. (2023). Pilot study of food-specific go/no-go training for overweight individuals: Brain imaging data suggest inhibition shapes food evaluation. *Social Cognitive and Affective Neuroscience*, *18*(1), Article nsab137. https://doi.org/10.1093/scan/nsab137

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2), 1–27. https://doi.org/10.1037/h0025848

Zoltak, M. J., Holland, R. W., Kukken, N., & Veling, H. (2020). Training choices toward low value options. *Judgment and Decision Making*, *15*(2), 254–265. https://doi.org/10.1017/S1930297500007397

Zoltak, M. J., Veling, H., Chen, Z., & Holland, R. W. (2018). Attention! Can choices for low value food over high value food be trained? *Appetite*, *124*, 124–132. https://doi.org/10.1016/j.appet.2017.06.010

Received April 20, 2024
Revision received November 29, 2024
Accepted January 10, 2025 ■