

# How choice confidence shapes preferences: a Registered Report

Isa Steinecker<sup>1,2</sup>, Philipp Sterzer<sup>1,2</sup>, Matthias Guggenmos<sup>1\*</sup>

<sup>1</sup>Visual Perception Laboratory, Department of Psychiatry and Psychotherapy, Charité –Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt Universität zu Berlin, and Berlin Institute of Health , 10117 Berlin, Germany

<sup>2</sup>Bernstein Center for Computational Neuroscience, Berlin, Germany

\* Corresponding author: Matthias Guggenmos ([matthias.guggenmos@charite.de](mailto:matthias.guggenmos@charite.de))

## **Abstract**

An intriguing finding in human decision making is that the act of choice itself affects preferences – chosen options often increase and unchosen options decrease in subjective value. The observation of such choice-induced revaluation inspired the theory of cognitive dissonance, which posits that human behavior generally strives for internal consistency. Yet, which aspects of choices are responsible for choice-induced revaluation has remained unclear. Here we argue that previous research has overlooked an important explanatory variable: choice confidence. While the original cognitive dissonance theory posits stronger revaluation for low-confidence decisions, a recently proposed theory that parallels confidence and reward predicts stronger revaluation for high-confidence decisions. To arbitrate between these hypotheses, and a confidence-independent mere-choice hypothesis, we employ an instrumental conditioning paradigm and quantitatively compare computational models of choice revaluation that operationalize the hypothesized confidence effects. Our results will clarify the mechanism of how choices and preferences interact in human decision making.

-

## Introduction

According to the standard view of human decision making, preferences determine choices, with an unidirectional arrow of causality. For instance, models of instrumental learning and decision making posit that preference changes are controlled by reward prediction errors (Rescorla and Wagner, 1972) and those preference changes subsequently influence choice behavior. Yet, longstanding research investigating preference dynamics in human decision also supports a reverse causal path, according to which the act of choice itself can affect preferences. Early work by Jack Brehm found that when participants made choices between pairs of objects, the very act of choice alone led to a preference shift: chosen objects increased in subjective value, unchosen objects decreased in subjective value (Brehm, 1956). Subsequent work replicated such post-decisional preferences changes – henceforth *choice revaluation* – in a large number of domains and paradigms in subsequent decades (e.g. Gerard & White, 1983; Greenwald, 1969; Shultz et al., 1999; Steele et al., 1993; White, 1981).

The findings by Brehm in particular, gave rise to the Leon Festinger’s now classic work “A Theory of Cognitive Dissonance” (Festinger, 1957), in which Festinger proposed that *cognitive dissonance* – internal psychological consistency or inconsistency – acts as a fundamental organizing force in the mind. Festinger reasoned that, when we are faced with difficult decisions, there will likely exist favourable aspects of rejected objects, thus inducing cognitive dissonance in the decision-maker. To reduce dissonance, we engage in a process of choice revaluation that might concern both the chosen and the unchosen option.

In 2008, a critical flaw in the logic of early paradigms was discovered, calling much of the research on choice revaluation into question (Chen, 2008; Chen and Risen, 2010). In earlier paradigms, researchers preselected choice items based on identical ratings in a prior preference rating phase. Subsequently, participants had to choose between the preselected items and rated them again, post-choice. The persistent finding was that the increase in ratings for chosen options (post- minus pre-choice) was higher than for unchosen options, which had since been taken as evidence for choice revaluation consistent with a reduction of cognitive dissonance. However, Chen and Risen demonstrated that this effect is to be expected for purely mathematical reasons: since pre-choice ratings have some degree of noise, participants might have a preference even between two identically rated items; conditioning on the chosen option therefore introduces a bias towards chosen options in the post-choice ratings.

Yet, subsequent research took this flaw into consideration and largely confirmed the effect of choice revaluation. Slight but clever modifications of the original paradigm (using a rating-rating-choice control condition; Coppin et al., 2010, 2012; Koster et al., 2015; Sharot et al., 2012), blind or fake choice paradigms (Sharot et al., 2010; Nakamura and Kawabata, 2013; Johansson et al., 2014; Luo and Yu, 2017) and evidence from neuroimaging studies (Sharot et al., 2009; Izuma et al., 2010, 2015;

Luettgau et al., 2020; Voigt et al., 2020) have recently provided new evidence and reinstated trust to the idea of choice revaluation.

What has remained unknown, is which aspects of the choice process are critical for choice revaluation. The proposal in this Registered Report is that *choice confidence* is a key characteristic of choices that may determine when – and how – choices affect preferences. Indeed, this idea is implicitly contained also in the theory of cognitive dissonance, which posits that choice revaluation is necessary precisely when choices are hard and thus choice confidence is low. Surprisingly, however, the effect of confidence on choice revaluation has not yet been systematically investigated. According to cognitive dissonance theory, the prediction for a role of confidence in choice revaluation is straightforward: lower confidence in decisions should lead to larger changes in subjective value. Henceforth, we refer to this as the *cognitive-dissonance hypothesis*.

On the other hand, recent findings in metacognition research make the case for another – and opposing – hypothesis about the role of confidence in choice revaluation. Growing evidence indicates that having confidence in a choice shares many of its affective, neurobiological and behavioral signatures with those of explicit reward feedback. At the behavioral level, two studies found that participants rated choices as more pleasing in which they were also more confident (Schwarze et al., 2013; Clos et al., 2015), suggesting that high-confidence responses are associated with higher subjective value. Consistent with this behavioural observation, a number of neuroimaging studies have found that, in the absence of external feedback, high- versus low-confidence choices activate brain regions of the reward network, in particular the ventral striatum and the ventral tegmental area (Daniel and Pollmann, 2010, 2012; Schwarze et al., 2013; Guggenmos et al., 2016; Hebart et al., 2016). Moreover, confidence prediction errors, i.e. the difference between expected and actual confidence, well explain the dynamics of learning without feedback in perceptual categorization (Daniel and Pollmann, 2012) and perceptual learning (Guggenmos et al., 2016) tasks. Together, these findings suggest that confidence acts as an internally-generated reinforcement signal when no external feedback is available. According to this notion, higher confidence is associated both with a stronger reinforcement signal and choice revaluation effects should thus become more pronounced with increasing choice confidence. We refer to this idea as the *confidence-learning hypothesis*.

In addition, it is possible that choice revaluation effects are independent of choice confidence. Indeed, it has been found that *the act of choice itself* is rewarding – regardless of the anticipated valence of choice outcomes or the difficulty of the choice. It has been suggested that the intrinsic rewarding nature of choice results from a sense of control or a sense of agency (see Leotti et al., 2010, for review). In line with this notion, neuroimaging studies have shown that the act of choosing activates reward-related brain areas irrespective of outcomes or motor actions (Sharot et al., 2009; Leotti and Delgado, 2011, 2014; Fujiwara et al., 2013). According to the *mere-choice hypothesis*, choice revaluation is thus independent of choice confidence.

To clarify the role of confidence in choice revaluation, we developed a paradigm based on instrumental conditioning that circumvents the flaw discovered by Chen and Risen (Chen, 2008; Chen and Risen, 2010). Each block of the experiment starts with an instrumental *acquisition phase*, in which participants learn the reward of six newly conditioned stimuli (CS), where two of each are of objectively identical monetary value. In the subsequent *revaluation phase*, participants are repeatedly presented with specific choices between two CS chosen in a way that, among each of the equal-reward CS pairs, one CS has a higher choice probability than the other. No reward feedback is provided in this phase. If choice revaluation occurs, the subjective value of the CS with higher choice probability should, on average, increase more than the value of the CS with lower choice probability. In the third and final *probe phase*, we put this to the test: for the first time, participants have to make choices between the equal-reward CS pairs and, in addition, rate all CS. A key aspect of this design is that choice revaluation effects are not only evaluated in terms of subjective ratings, but also in terms of the effective consequences on instrumental behavior. Moreover, since the analytic rationale of the present study is based on objectively equal rewards, and not subjective ratings, it is immune against the aforementioned design flaw (Chen, 2008; Chen and Risen, 2010). Mathematically, no difference in value ratings or choice preferences can be expected between two CS associated with equal objective reward.

To arbitrate between the three hypotheses - *cognitive-dissonance*, *confidence-learning*, and *mere-choice hypothesis* - about the role of confidence in choice revaluation, we will use computational models that leverage the framework provided by reinforcement learning models. Current models of reinforcement learning assume a forward link between outcomes, preferences and choices: outcomes shape preferences by means of reward prediction errors, that is, the difference between expected and obtained rewards (Rescorla & Wagner, 1972), and those updated preferences influence future choice behavior. They do not, however, consider that the act of a choice itself might affect preferences in the absence of any outcomes. It is also this gap in current models of reinforcement modeling that motivates the present work. In line with our three hypotheses about the role of confidence in choice revaluation, we devised models that are based on regular reinforcement learning models when external feedback is available, but are augmented with a choice revaluation term in the absence of feedback.

Overall, our research questions and associated analyses will be as follows (see also the Design Table). We will evaluate whether choice revaluation affects subjective value ratings (Confirmatory planned analysis 1) and subsequent choice behavior (Confirmatory planned analysis 2) in instrumental conditioning, consistent with a preference increase for the chosen option. While these first two analyses serve to assess the basic behavioral effect of choice revaluation, our key analysis is based on a formal model comparison among model variants that differ in terms of their implementation of choice revaluation (Confirmatory planned analysis 3). In particular, we will compare three models that test our three alternative hypotheses about the role of confidence in choice revaluation.

Additional planned but non-confirmatory analyses will evaluate how sustained choice revaluation effects change across the time course of the experiment, and whether there is evidence for self-reinforcement in phases without external feedback, in line with the confidence-learning hypothesis.

## Methods

### Participants and online experiment

The present study is set up as an online experiment that runs in web browsers. 185 participants will be recruited through the online recruiting platform Prolific (<http://prolific.co>). As the large majority of registered participants at Prolific are native English speakers, all instructions and other texts in the experiment will be written in English and being a native English speaker is therefore an inclusion criterion. Apart from the language requirement and a minimum age of 18, no other a priori inclusion or exclusion criteria will be applied. At the beginning of the experiment, participants can select a checkbox stating “I will not cheat in this experiment”. However, although not explicitly specified, this checkbox is optional (Clifford and Jerit, 2016). The experiment itself is programmed in JavaScript, using the library p5.js (<http://p5.js.org>) and JATOS as a hosting platform (Lange et al., 2015; <http://www.jatos.org/>).

A posteriori exclusion criteria will be low performance (average performance < 60% correct in the acquisition and revaluation phase combined), incongruent confidence ratings (confidence higher for incorrect than correct choices), too long inter-trial intervals (ITI) which may indicate cheating (average ITI > 10 secs), and too many attempts in the match-up memory game which may indicate either inattentiveness or memory deficits (average number of attempts > 16).

The sample size calculation is based on the two main behavioral effects underlying the general hypothesis of choice revaluation in instrumental conditioning, i.e. that choices affect subjective value ratings (Confirmatory planned analyses 1) and instrumental decision making (Confirmatory planned analyses 2). It is expected that the latter effect will be smaller for two reasons: 1) choices are binary in the present study and are thus much more noisy than graded value ratings, 2) compared to preferences, effects on choices are one step further in the causal chain and thus are likely to come with additional sources of cognitive noise. In the absence of prior estimates for the effect size of choice revaluation on instrumental decision making, a minimal effect of interest of 3% is defined, that is, CS of set 2 are chosen 3% more often than expected by chance (see Experimental procedures for explanation). With 18 samples per participant (6 blocks and 3 critical choices per block), the required sample size to detect the minimal effect at 0.95 a priori power and using a one-tailed paired t-test is N=168. We expect a dropout of approximately 10% and will resample until the desired sample size is achieved.

The present study was approved by the ethics committee of the Charité Universitätsmedizin Berlin. Informed consent will be obtained from all participants prior to the start of the experiment.

## Experimental procedures

### *Stimuli*

CS will be abstract colored symbols chosen from an online image database of Mayan and Aztek symbols (N=79), consistent with the cover story of the experiment (see section Cover story and instructions). The symbols have roughly identical sizes of around 400x400 pixels. For each participant, 36 stimuli will be randomly drawn from the stimulus pool, corresponding to the 6 CS presented in each of the 6 experimental blocks (see Experimental design). Stimulus examples are shown in Figure 1.

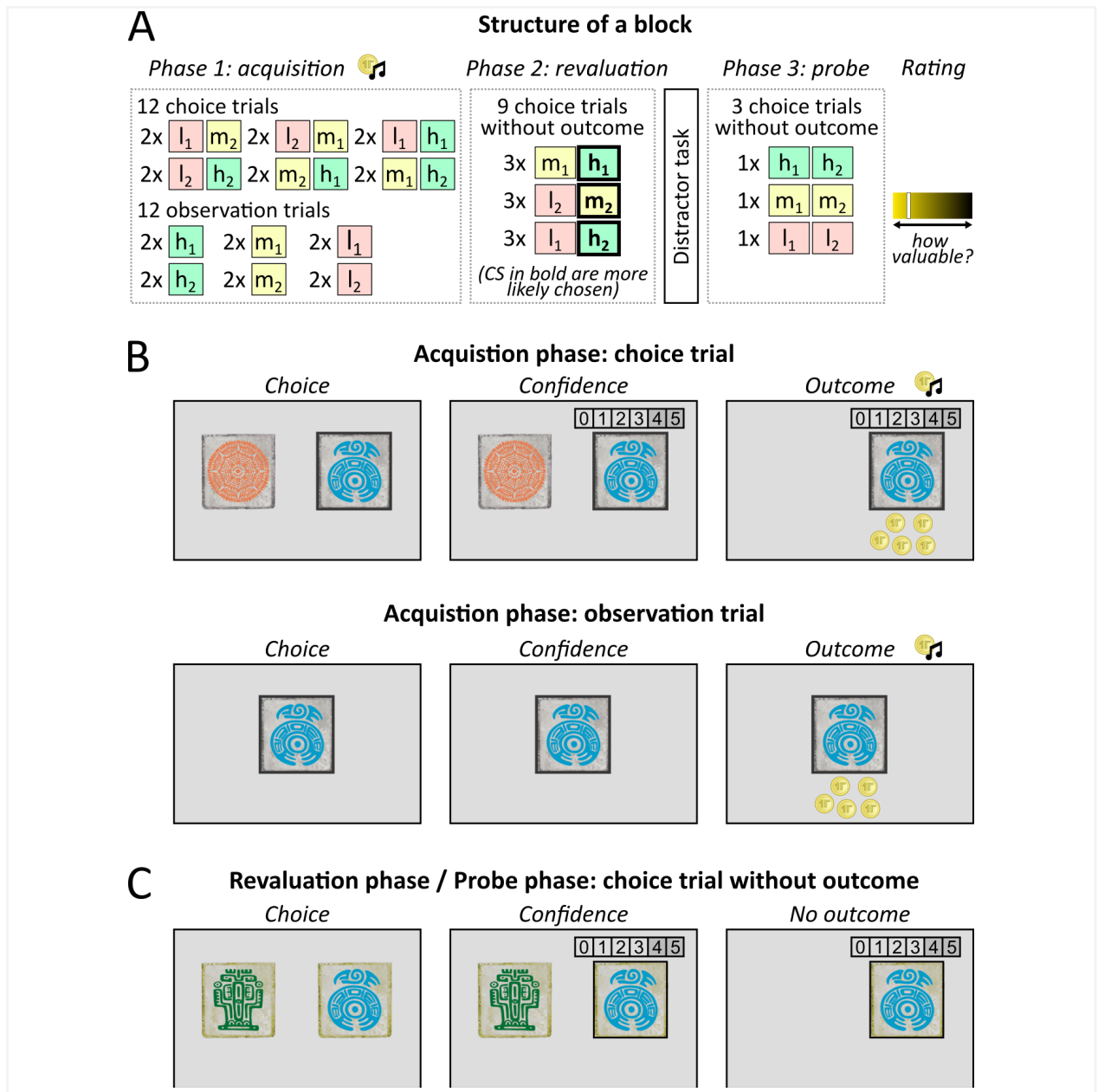
### *Experimental design*

To probe choice revaluation in instrumental decision making, a paradigm was devised that closely resembles conventional value-based decision making tasks but includes a no-feedback phase to isolate choice-induced from feedback-induced value changes. All design contrasts are within-subject.

The experiment comprises six blocks with an identical structure (Figure 1A), in each of which participants have to learn about the reward of six CS. The six CS are assigned objective rewards such that two CS have identical low rewards (CS  $l_1$  and CS  $l_2$ ), two CS have identical medium rewards (CS  $m_1$  and CS  $m_2$ ) and two CS have identical high rewards (CS  $h_1$  and CS  $h_2$ ). For reasons explained below, the CS  $l_1/m_1/h_1$  are referred to as *set 1* and  $l_2/m_2/h_2$  as *set 2*.

A block starts with an *acquisition phase* (24 trials) in which participants receive trial-wise feedback to learn about the rewards of the CS, either in the form of choice trials (50%) or observation trials (50%) (Figure 1B). In choice trials, participants have to select one of two CS presented on the left and right side of the screen center by means of a mouse click. After the mouse click, a confidence scale appears above the chosen CS on which participants have to indicate their confidence in the choice on a discrete 6-point scale, where 0 corresponds to “guessing” and 5 to “100% sure about the choice”. The selection of the confidence rating is likewise performed by mouse. Following the confidence rating, the unchosen CS disappears and monetary outcome is presented in the form of coins below the chosen CS. Note that the choice pairs in the acquisition phase do neither include choices between equal-reward CS, nor choices that are presented in the revaluation phase (see next paragraph). Moreover, the choice pairs are matched between sets 1 and 2, such that the two CS of each reward level are paired with CS of identical reward levels. Observation trials are included to enforce a more balanced learning of all CS. Here, only one CS is presented centrally on the screen. Participants are instructed to click on the CS to see and obtain the outcome for the CS.





**Figure 1. Design and trial structure.** **A)** Each block consists of three phases. In the *acquisition phase*, observers learn the rewards of overall six CS where two CS each are assigned to a low-reward category ( $l_1/l_2$ ), a medium-reward category ( $m_1/m_2$ ) and a high-reward category ( $h_1/h_2$ ). The *revaluation phase* includes choice pairs that induce an asymmetry between the two CS of each reward category, such that CS of set 2 ( $l_2/m_2/h_2$ ) are more likely to be chosen than CS of set 1 ( $l_1/m_1/h_1$ ). In the *probe phase*, which follows after a brief distractor task (a match up memory task), observers are subjected to choices between each of the two equal-reward CS, which allows testing the hypothesis of stronger choice revaluation for CS of set 2 relative to CS of set 1. Note that feedback is provided only in the acquisition phase. At the end of the block, observers rate how valuable each CS feels to them. **B)** Trial structure in the acquisition phase. In choice trials, participants have to make a choice between two CS and indicate the confidence in their choice. They receive rewards according to a fixed reward schedule. The acquisition phase also includes observation trials to achieve a more balanced learning of all CS. Here, participants click on a single CS to obtain the reward. **C)** Trial structure in the revaluation and probe phase. The revaluation and probe phase consists of choice trials which are identical to those in the acquisition phase, except that no outcome is shown. The stone switches underneath the symbols are shaded slightly golden to indicate that these are choices without outcomes.

The critical intervention is the subsequent *reevaluation phase*, in which participants are repeatedly presented with three specific choice pairs designed to induce an asymmetry in the choice probability between each of the two equal-reward CS. The three choice pairs are selected such that within each of the two equal-reward CS, the CS of set 2 has higher choice probability than the CS of set 1:  $(m_1, h_1)$ ,  $(l_2, m_2)$ ,  $(l_1, h_2)$ . These pairs are referred to as *reevaluation pairs*. For instance,  $h_1$  is paired with a medium-reward CS ( $m_1$ ), while  $h_2$  is paired with a low-reward CS ( $l_1$ ). Due to the expected imbalance in value differences, participants will more often choose  $h_2$  compared to  $h_1$ . As a consequence, a positive choice revaluation should, on average, be greater for  $h_2$  relative to  $h_1$ . The same logic applies to the CS of the other two reward levels, i.e. the CS of set 2 is always assigned a higher objective choice probability compared to the CS of set 1 (a similar design was recently used in the context of classical conditioning; Luettgau et al., 2020). Note that the revaluation pairs are not presented in the acquisition phase, hence the revaluation phase is the first time for participants to choose between them. The revaluation phase consists of three trial sequences, in each of which the three revaluation pairs are presented once in random order ( $3 \times 3 = 9$  trials). To avoid a mix of feedback-induced and choice-induced value changes, participants do not receive feedback in the revaluation phase (Figure 1C). However, participants are instructed that they eventually see and receive the outcome of these choices after each block, i.e. the choices are incentivised (see Cover story).

After a brief distraction with a memory task (an unrelated match-up memory task taking approximately 1 minute), choice revaluation effects are probed in the third and final phase, the *probe phase*. To test whether the revaluation phase induces differences between CS associated with objectively equal rewards, participants have to choose once between each of the three equal-reward pairs  $(l_1, l_2)$ ,  $(m_1, m_2)$ ,  $(h_1, h_2)$ , in randomized order. As in the revaluation phase, no outcomes are shown in the probe phase (Figure 1C). Immediately after these three choices, participants have to indicate on a continuous scale how valuable each of the six CS feels to them. After the last rating, participants are revealed the cumulative outcome of their choices during the revaluation and the probe phase. A block ends with a display of a participant's performance for the entire block, computed as the percentage choices in which the CS associated with greater or equal reward was chosen.

Outcomes are animated as coins rapidly appearing one after another below a CS. Coins in the acquisition phase are labeled with "1 $\Gamma$ ", where  $\Gamma$  is the currency of the game (see Cover story). The positions of the coins are pseudo-randomized (non-overlapping) and give the impression of coins falling on a table. The visual presentation of coins is accompanied by a glittering sound which lasts as long as the coin animation (i.e., lasting longer when more coins are obtained). While the number of coins obtained for a cue is exactly the same for each choice of a particular CS, the random positioning and the relatively large number of coins (see next paragraph) introduces some uncertainty about the exact value of a CS. The purpose is for participants to develop a *feeling* for the

value of each CS rather than putting an *explicit number* on each symbol which would make too easily apparent the ambiguity of choices in the probe phase. For similar purposes, previous studies have often drawn rewards from a reward distribution rather than setting a fixed reward. However, here this method is not used, as it comes with the risk that average rewards are not actually identical between two equal-reward CS, especially at low trial numbers such as in the present study.

The rewards associated with the CS were chosen based on four considerations. First, rewards should be reasonably high to make counting of the coins impossible (note the number of coins presented corresponds to the reward in  $\Gamma$ ). Second, reward differences should increase exponentially with the absolute number of coins, accounting for the logarithmic nature of counting precision (Kaufman et al., 1949). Based on pilot data, a set of 10 exponentially increasing rewards was derived: 12 $\Gamma$ , 16 $\Gamma$ , 20 $\Gamma$ , 25 $\Gamma$ , 30 $\Gamma$ , 36 $\Gamma$ , 42 $\Gamma$ , 49 $\Gamma$ , 56 $\Gamma$ , 64 $\Gamma$ . For simplicity, these rewards are referred to as reward levels 1-10. Third, there should be variance in CS rewards between blocks to reduce learning transfer between blocks. Therefore, in half of the blocks, the low-reward CS ( $l_1/l_2$ ) are assigned to reward level 1 and the high-reward CS ( $h_1/h_2$ ) to reward level 7 (“low overall reward level”), whereas in the other half, low-reward CS are assigned to reward level 4 and high-reward CS to reward level 10 (“high overall reward level”). The rewards of medium-reward CS ( $m_1/m_2$ ) are inbetween. Fourth, to introduce variance in the reward differences of the revaluation phase, the rewards for medium-reward CS ( $m_1/m_2$ ) are chosen in a way to systematically vary the reward differences of the choice revaluation pairs ( $m_1, h_1$ ) and ( $l_2, m_2$ ). Table 1 provides an overview over all CS rewards and the reward differences in the revaluation phase.

**Table 1. Reward schemes for the six experimental blocks and reward differences in the revaluation phase.** Each row of the table indicates the reward scheme of one block (6 blocks in total). The order is randomized for each participant. CS reward levels and rewards (grey background): CS reward levels are coded from 1 to 10. The sub-columns  $l_1/l_2$ ,  $m_1/m_2$  and  $h_1/h_2$  display the reward levels assigned to low-, medium- and high-reward CS, and in brackets their corresponding rewards in the currency of the experiment ( $\Gamma$ ).

Reward differences in the revaluation phase (white background): The reward schemes are defined with the goal to introduce an asymmetry between the equal-reward CS  $l_1/l_2$ ,  $m_1/m_2$  and  $h_1/h_2$ . The asymmetry is introduced in the revaluation phase, in which the reward differences of the choice pairs ( $l_1, h_2$ ), ( $l_2, m_2$ ), ( $m_1, h_1$ ) are defined such that between each of the two equal-reward CS, the CS of set 2 ( $l_2/m_2/h_2$ ) is more likely chosen than the CS of set 1 ( $l_1/m_1/h_1$ ).

	Reward levels (brackets: reward in the experimental currency $\Gamma$ )			$l_1/l_2$ : reward differences in the revaluation phase		$m_1/m_2$ : reward differences in the revaluation phase		$h_1/h_2$ : reward differences in the revaluation phase	
	$l_1/l_2$	$m_1/m_2$	$h_1/h_2$	set 1: $l_1-h_2$	set 2: $l_2-m_2$	set 1: $m_1-h_1$	set 2: $m_2-l_2$	set 1: $h_1-m_1$	set 2: $h_2-l_1$
1	1 (12 $\Gamma$ )	2 (16 $\Gamma$ )	7 (42 $\Gamma$ )	-6	-1	-5	+1	+5	+6
2	1 (12 $\Gamma$ )	4 (25 $\Gamma$ )	7 (42 $\Gamma$ )	-6	-3	-3	+3	+3	+6
3	1 (12 $\Gamma$ )	6 (36 $\Gamma$ )	7 (42 $\Gamma$ )	-6	-5	-1	+5	+1	+6
4	4 (25 $\Gamma$ )	5 (30 $\Gamma$ )	10 (64 $\Gamma$ )	-6	-1	-5	+1	+5	+6
5	4 (25 $\Gamma$ )	7 (42 $\Gamma$ )	10 (64 $\Gamma$ )	-6	-3	-3	+3	+3	+6
6	4 (25 $\Gamma$ )	9 (56 $\Gamma$ )	10 (64 $\Gamma$ )	-6	-5	-1	+5	+1	+6

After the third and the sixth block, participants are again presented with choices between the CS of equal-reward pairs. The primary purpose of this phase is to break the uniform structure of the experiment and to thereby increase motivation. At the same time, the responses in this phase allow assessing whether choice biases are present after a delay much longer compared to the probe phase.

This phase is therefore referred to as the *delayed probe phase*. It should be noted that choices in this phase may be subject to a consistency bias such that participants try to respond in line with their choices and/or ratings in the regular first probe phase. In addition, participants will have seen outcomes for their choice in the corresponding first revaluation and probe phase (although in a cumulative manner, which should make it impossible for them to use this information for stimulus-specific value updates). Due to these two potential issues, the analysis of the delayed probe phase is not considered confirmatory.

### *Cover story and instructions*

To increase task engagement and to motivate, in particular, the occurrence of choices without feedback, the experiment will be wrapped in a cover story and includes elements of gamification (Hamari et al., 2014). Participants are told to imagine themselves as part of a group of explorers who just discovered an ancient temple filled with treasures on several floors (each level of the temple corresponds to an experimental block). The treasures are stored in coin vaults located invisibly behind walls. The explorer repeatedly encounters pairs of switches (or occasionally single switches in case of observation trials), which give access to these coin vaults. Switches are labeled with symbols which provide information about the value of the coin vault connected to the switch. Importantly, pressing one of the two switches blocks the other, thus the explorer has to make a choice between the two. The task is therefore to “crack the code” of the symbols, i.e. learn about the value of the symbols (acquisition phase). Since new symbols appear in each level, learning starts anew when participants arrive at the next level of the temple.

At the end of each level, the explorer encounters the “King’s chamber” (choice revaluation and probe phase). This chamber is said to have two special characteristics. First, the explorer does not see the outcomes directly after each choice; instead the entire treasure collected in the King’s chamber is released with a special switch at the far end of the chamber. Second, for the coin vaults of the chamber, the temple is said to be generous and triples the amount of earned coins (to alleviate a potential decrease of motivation due to the absence of feedback). Close to the end of the chamber (i.e., end of the revaluation phase), the explorer encounters a locked door which must be unlocked by solving a brief match-up memory game (distractor task). Once the memory is solved, the door unlocks and the explorer sees the remaining switch pairs of the chamber (probe phase). At the end of the level, one’s companions ask how valuable each of the symbols feels (rating) and then the switch appears, which releases all coins collected in the chamber.

The coins released by the coin vaults are of the ancient currency “ $\Gamma$ ” and the conversion rate to participants’ preferred real currency is revealed at the end of the experiment. Coins are labeled with “1 $\Gamma$ ”. Participants are instructed that they will be paid off all coins collected during the experiment, with the specified conversion rate between  $\Gamma$  and their preferred currency. See supplementary section “Instructions” for the verbatim instructions provided to participants.

## Computational modeling

The goal of computational modeling will be twofold. The first goal will be to test whether there are detectable dynamics of choice revaluation, such that they can be quantitatively described by a mechanistic model. The second goal will be to differentiate between different hypotheses about a possible role of confidence in choice revaluation.

Computational modeling will proceed in two stages. Stage 1 applies to the data of the acquisition phase, for which a conventional reinforcement learning model with a Rescorla-Wagner learning rule will be employed. Accordingly, subjective values will be updated in proportion to the difference between expected values and observed values:

$$v_c \rightarrow v_c + \alpha(b, t) (r - v_c) \quad (1)$$

where  $v_c$  is the value estimate of the chosen option,  $r$  is the obtained reward and  $\alpha$  is the learning rate. We allow the learning rate to vary as a function of block  $b$  and trial  $t$  as follows:

$$\alpha(b, t) = (\alpha_0 + \alpha_1 b) e^{-\alpha_2 t} \quad (2)$$

We anticipate that learning rates reduce within a block, as the learning process stabilizes, and increase across blocks, as participants become more familiar with the task.

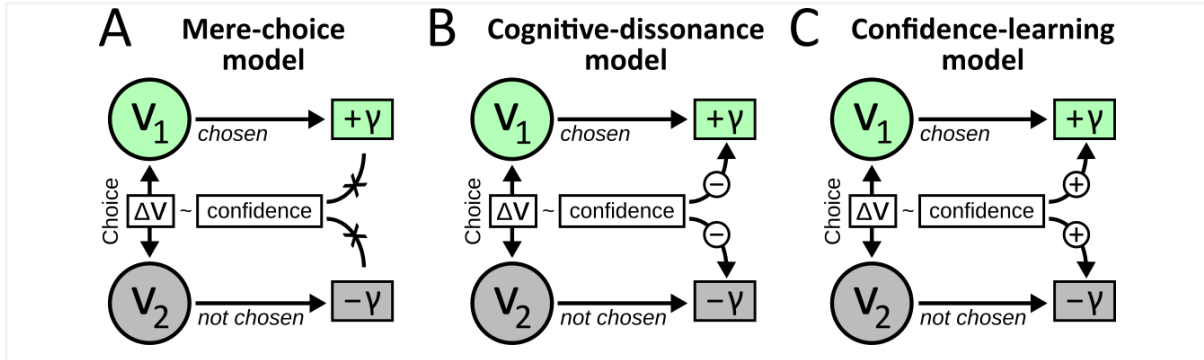
In each trial, choices between two options with values  $v_1$  and  $v_2$  will follow a softmax action selection rule:

$$p(v_1) = \frac{1}{1 + \exp(-\beta(v_1 - v_2))} \quad (3)$$

where  $p(v_1)$  is the probability to choose option  $v_1$  and  $\beta$  is the inverse decision noise temperature.

A novel aspect of our modeling procedure will be an inference step about the *effective rewards* of the rewards presented, since it cannot be assumed that if  $N$  coins are presented to the participant this is indeed perceived as a reward of  $N$ . Therefore, the effective reward of the low-reward CS ( $I_1 / I_2$ ) will be fixed to  $r_l = 1$  and two parameters  $r_m$  and  $r_h$  will be introduced which represent the effective rewards associated with medium-reward and high-reward CS with the constraint  $r_l < r_m < r_h$ .

Stage 2 corresponds to the phases without feedback (revaluation and probe phase) in which subjective values may or may not be influenced by choice revaluation effects. Note that external feedback is not provided in the revaluation and probe phase. Models will be initialized with the final values  $v_i$  of Stage 1. Three model families will be tested with choice revaluation terms and which differ with respect to the role of choice confidence (Figure 2).



**Figure 2. Computational models.** The depicted example shows a choice between two CS with subjective values  $v_1$  and  $v_2$ . The CS colored in green with subjective value  $v_1$  is chosen. The confidence in the choice is assumed to depend on the absolute value difference  $\Delta v = |v_1 - v_2|$ . Three types of models will be tested which make different predictions regarding whether and how confidence modulates choice revaluation. **A)** According to the mere-choice model, it is the act of the choice itself irrespective of choice confidence that affects values. Values are updated by a constant  $\gamma$ . **B)** According to the cognitive dissonance theory, choice-induced value updates reflect dissonance resolution, predicting stronger effects for harder choices (i.e. smaller  $\Delta v$ ). Confidence thus acts as a negative modulator of choice revaluation. **C)** The confidence-learning model considers confidence as a reinforcement learning signal, predicting larger learning effects for higher confidence. Confidence thus acts as a positive modulator of choice revaluation.

A *mere-choice model* assumes that the act of the choice itself – irrespective of the certainty in that choice – is responsible for choice revaluation. For this model, subjective values  $v$  will thus be updated by the *revaluation learning rate*  $\gamma$ , independent of choice confidence:

$$\text{Mere-choice model: } v \rightarrow v \pm \gamma \quad (4)$$

The  $\pm$  sign indicates that the learning rule either increases (chosen option) or decreases (unchosen option) the value.

A *cognitive dissonance model* assumes that value changes reflect a dissonance resolution. Harder choices, i.e. choices between CS closer in value and about which observers are hence less confident, require a larger amount of dissonance resolution. Confidence – operationalized as the normalized value difference, thus negatively modulates choice revaluation:

$$\text{Cognitive-dissonance model: } v \rightarrow v \pm \gamma \left[ 1 - \frac{|v_c - v_u|}{v_c + v_u} \right] \quad (5)$$

The normalisation by the sum of both values ensures that the confidence term  $1 - |v_c - v_u| / (v_c + v_u)$  approaches 0 for large value differences and 1 for small values differences.

Finally, a *confidence-learning* model regards confidence as a learning signal, such that confidence positively modulates choice revaluation:

$$\text{Confidence-learning model: } v \rightarrow v \pm \gamma \frac{|v_c - v_u|}{v_c + v_u} \quad (6)$$

Here, the confidence term  $|v_c - v_u| / (v_c + v_u)$  approaches 1 for large value differences and 0 for small values differences.

Not all modeling decisions that are implicit to the above equations can be a priori fixed based on current knowledge, which poses a challenge within the Registered Report framework. For instance, previous evidence of whether choice revaluation affects only the chosen option, or both the chosen and the unchosen option has been mixed, thus giving no justification to specify this model part a priori. As we are nevertheless convinced that the advantages of Registered Reports outweigh these challenges, for the present work we decided to preregister and expose those specific modeling decisions that will be decided *conditional on the data*. Specifically, each of the three model families is tested in the following additional variants:

- i. In which learning is proportional to the absolute value  $v$  (i.e. the term  $\gamma$  becomes  $\gamma v$ )
- ii. In which the confidence term has an additional slope parameter (i.e. the term  $|v_c - v_u| / (v_c + v_u)$  becomes  $|v_c - v_u| / (v_c + v_u + \vartheta)$ )
- iii. With different revaluation learning rates for the chosen and unchosen option (i.e. introducing two update parameters  $\gamma_c$  and  $\gamma_u$ )
- iv. Without an update of the unchosen option (i.e., models are reduced to an update of the chosen option only)

This creates a model space of overall 48 models (3 families  $\times$   $2^4$  variants). In addition, a null model will be tested according to which no choice revaluation effects occur in phases without feedback, i.e. value learning stops after the acquisition phase.

To fit model parameters, we will use a hierarchical Bayesian modeling approach (van Geen and Gerraty, 2020). A key advantage of this approach over conventional maximum likelihood estimation on the single-subject level is that parameter estimates are regularized by an empirical prior derived from group-level information. In brief, the idea is to enforce a parametric distribution of parameters across subjects, whereby the distribution parameters are themselves estimated from the data. As demonstrated by van Geen and Gerraty (2020), this hierarchical approach is both more data-efficient (as empirical priors are estimated from within the data and do not require independent studies) and better in terms of out-of-sample prediction. Table 2 provides an overview of the distributional assumptions of all modeling parameters, which are largely based on recommendations by Gershman (2016) and van Geen and Gerraty (2020).

We will use the Python package PyMC3 for Bayesian modeling (Salvatier et al., 2016). The No-U-Turn-Sampler will be used to sample from the posterior distribution (Hoffman and Gelman, 2014).

**Table 2.** Prior distributions for subject-level parameters and hyperparameters at the group level.

Parameter	Distribution	level
Reward learning rate: intercept	$\alpha_0 \sim \text{Beta}(\text{shape}_1, \text{shape}_2)$	subject
	$\text{shape}_1 \sim \text{Cauchy}^+(\text{loc} = 0, \text{scale} = 25)$	group
	$\text{shape}_2 \sim \text{Cauchy}^+(\text{loc} = 0, \text{scale} = 25)$	group
Reward learning rate: linear slope across blocks	$\alpha_1 \sim \text{Normal}(\text{loc} = 0, \text{scale} = 1)$	subject
Reward learning rate: exponential decay across trials	$\alpha_2 \sim \text{Normal}^+(\text{loc} = 0, \text{scale} = 1)$	subject
Decision noise	$\beta \sim \text{Gamma}(\text{shape}, \text{scale})$	subject
	$\text{shape} \sim \text{Cauchy}^+(\text{loc} = 0, \text{scale} = 25)$	group
	$\text{scale} \sim \text{Cauchy}^+(\text{loc} = 0, \text{scale} = 25)$	group
Revaluation learning rate	$\gamma \sim \text{Beta}(\text{shape}_1, \text{shape}_2)$	subject
	$\text{shape}_1 \sim \text{Cauchy}^+(\text{loc} = 0, \text{scale} = 25)$	group
	$\text{shape}_2 \sim \text{Cauchy}^+(\text{loc} = 0, \text{scale} = 25)$	group
Confidence slope parameter	$\vartheta \sim \text{Normal}(\text{loc} = 0, \text{scale} = 10)$	subject
Effective rewards	$r_m \sim \text{Cauchy}^+(\text{loc} = r_l = 1, \text{scale} = 25)$	group
	$r_h \sim \text{Cauchy}^+(\text{loc} = r_m, \text{scale} = 25)$	group

## Confirmatory planned analyses

Confirmatory planned analyses are pre-specified and will serve to address the main hypotheses of this study.

Statistical analyses not involving computational modeling will be performed with linear mixed-effects models, thereby leveraging the ability to consider the hierarchical variance structure of our data (Magezi, 2015). The MixedLM implementation of the Python package statsmodels will be used (Skipper and Perktold, 2010). In the models, the variable *subject* will be considered a random effect and (if applicable) *block* a nested (to subject) random effect. All dependent and independent variables will be standardized prior to model fitting. Table 3 provides an overview and description of all variables that will be used in both the confirmatory and exploratory analyses.



**Table 3.** Overview of all variables that will be in the linear mixed-effects models, as well as descriptions and value ranges for each variable. Note that all variables are standardized prior to model fitting.

Variable	Description
<b>Fixed effects: dependent variables</b>	
<i>choice_side</i>	Whether the CS on the left ( <i>choice</i> =0) or the right side ( <i>choice</i> = 1) of the screen has been chosen
<i>confidence</i>	Choice confidence (0-5)
<i>correct</i>	Whether a response was incorrect ( <i>correct</i> = 0) or correct ( <i>correct</i> = 1)
<i>rating</i>	Subjective value rating of a CS; a continuous number between 0 and 1
<b>Fixed effects: independent variables of interest</b>	
<i>block_delayed</i>	Number of the block in which a choice pair in the delayed probe phase has previously appeared (1-6)
<i>condition</i>	Whether a CS belongs to set 1 ( <i>condition</i> = 0) or set 2 ( <i>condition</i> = 1)
<i>condition_side</i>	Whether the CS on the right side of the screen belongs to set 1 ( <i>condition_side</i> = 0) or set 2 ( <i>condition_side</i> = 1)
<i>nchoices</i>	How often the current choice has occurred for the current choice pair in the revaluation phase (includes the current choice)
<i>trial</i>	Trial number within a block
<b>Fixed effects: independent control variables</b>	
<i>block</i>	Block (1-6); note that <i>block</i> is also used as a random effect
<i>decision_noise</i>	Decision noise parameter $\beta$ (derived from model fitting)
<i>delayed_probe_phase</i>	Whether a trial belongs to the first ( <i>delayed_probe_phase</i> = 0) or second ( <i>delayed_probe_phase</i> = 1) delayed probe phase
<i>reward</i>	Effective reward associated with a CS, as derived from the modeling parameters $r_l$ , $r_m$ , $r_h$
<i>reward_category</i>	Categorical predictor coding whether a CS belongs to the low-reward (= 0), medium-reward (= 1) or high-reward (= 2) category
<i>reward_diff</i>	Unsigned effective reward difference between two CS; see <i>reward</i>
<i>reward_diff_chosen</i>	Signed difference of effective rewards (chosen CS minus unchosen CS); see <i>reward</i>
<i>reward_diff_side</i>	Signed difference of effective rewards (CS on the right of fixation minus CS on left); see <i>reward</i>
<i>reward_level</i>	Whether a block is assigned to the low (= 0) or high (= 1) overall reward level
<i>reward_sum</i>	Sum of effective rewards of a choice pair; see <i>reward</i>
<i>value</i>	Subjective value of a CS; subjective values are obtained from model fits to the choice data of the acquisition phase (see Section Computational modeling)
<i>value_chosen</i>	Subjective value of the chosen CS; see <i>value</i>
<i>value_diff_chosen</i>	Signed subjective value difference (chosen CS minus unchosen CS); see <i>value</i>
<i>value_diff_side</i>	Signed subjective value difference (CS on the right of fixation minus CS on left); see <i>value</i>
<b>Random effects variables</b>	
<i>block</i>	Random effects predictor for block (1-6); note that <i>block</i> is also used as a fixed effect
<i>stimulus</i>	Random effects predictor for the specific CS presented (relevant for analyses of value ratings)
<i>subject</i>	Random effects predictor for subjects

### Confirmatory planned analysis 1: choices affect subjective value ratings

To test whether the mere act of choice increases the subjective value ratings of the chosen CS and decreases subjective value ratings of the unchosen CS, a linear mixed-effects analysis will be performed. The dependent variable *rating* reflects the subjective value rating at the end of each block. The key predictor to test the hypothesis will be *condition*, which codes whether the CS is of set

1 (lower choice probability during the revaluation phase) or set 2 (higher choice probability during the revaluation phase). In Wilkinson notation, the model will be specified as follows:

$$rating \sim condition * (value + reward + reward\_category + reward\_level) + block + decision\_noise + RE$$

The random effects term *RE* comprises random effects for *subject* (random intercept and random slopes for *condition* and *value*), *block* (random intercept; *block* is nested to *subject*) and *stimulus* (random intercept). The null hypothesis will be rejected if the predictor *condition* is significant ( $p < 0.05$ ; one-tailed).

#### *Confirmatory planned analysis 2: choices affect instrumental decision making*

To test whether hypothesized choice-induced changes in subjective value translate to instrumental decision making, choices of the probe phase will be analysed with a mixed regression model. The dependent variable *choice\_side* codes whether the CS on the right side (*choice\_side*=1) or left side (*choice\_side*=0) is chosen. Correspondingly, the key predictor *condition\_side* is 1 when the CS on the right side is of set 2, and 0 if the cue on the left side is of set 2 (the contralateral CS is always of set 1). In Wilkinson notation, the model will be specified as follows:

$$choice\_side \sim condition\_side * (value\_diff\_side) + condition\_side : (reward + reward\_category + reward\_level + block) + RE$$

The random effects term *RE* comprises random effects for *subject* (random intercept and random slopes for *condition\_side* and *value\_diff\_side*) and *block* (random intercept; *block* is nested to *subject*). The null hypothesis will be rejected if the predictor *condition\_side* is significant ( $p < 0.05$ ; one-tailed).

#### *Confirmatory planned analysis 3: formal model comparison*

Central to the present work is a formal model comparison (see section “Computational modeling” for an overview of models) for the probe phase. The model comparison will be based on the Watanabe-Akaike Information Criterion (WAIC; Watanabe, 2010). In a first step, the winning model within each model family (*mere-choice*, *cognitive dissonance*, *confidence learning*) is determined based on the WAIC. In a second step, the uncertainty (95% confidence interval) of WAIC is considered to statistically compare the winning models of each family as well as the null model, according to which no choice revaluation dynamics occur in the revaluation phase. The prediction is that choice revaluation exists (i.e. the null model is rejected); however none of the three choice revaluation hypotheses is favoured a priori. A model is declared superior to another, if the respective 95% WAIC confidence intervals are not overlapping.

### **Additional planned analyses**

Additional planned analyses are pre-specified, but serve either as sanity checks or elucidate aspects not central to the main hypotheses of this work.

### *Learning curves in the acquisition phase*

The purpose of the acquisition phase is for participants to learn about the reward values of the CS. To validate this basic requirement, a logistic mixed model will be fitted to test for a main effect of *trial* on responses *correct*. In Wilkinson notation, the model will be specified as follows:

$$correct \sim trial * (reward\_diff + reward\_sum + reward\_level + block) + decision\_noise + RE$$

The random effects term *RE* comprises random effects for *subject* (random intercept and random slopes for *trial* and *reward\_diff*) and *block* (random intercept; *block* is nested to *subject*). We expect a statistically highly significant main effect of *trial* on responses *correct*.

Likewise, participants' confidence should increase across the acquisition phase. To validate, a linear mixed-effects model will be fitted testing for a main effect of *trial* on responses *correct*. In Wilkinson notation, the model will be specified as follows:

$$confidence \sim trial * (reward\_diff + reward\_sum + reward\_level + block) + decision\_noise + RE$$

The random effects term *RE* comprises random effects for *subject* (random intercept and random slope for *trial* and *reward\_diff*) and *block* (random intercept; *block* is nested to *subject*). We expect a statistically highly significant main effect of *trial* on *confidence*.

### *Choice asymmetry in the revaluation phase*

As a validation of the paradigm, it will be assessed whether participants preferred CS of set 2 over CS of set 1 in the revaluation phase. For both sets, the following quantities will be reported: the average of the 1) proportion of choices, 2) the confidence in choices and 3) the reaction times for choices. The averages of both sets will be statistically compared by means of a two-sample proportions z-test (proportion of choices) or a t-test (confidence, logarithmized reaction times). We expect statistically highly significant effects in the predicted direction for all three dependent variables (proportion of choices, confidence, reaction times).

### *Persistence of choice revaluation effects*

If choice revaluation effects are observed in the probe phase, the question arises how sustained such revaluation effects are and whether they can be observed at later points of the experiment. To this aim, trials of the two delayed probe phases (after the third and sixth block) will be analysed, in which observers are again presented with choices between all equal-reward CS from the preceding three blocks. The delay in delayed probe phases is thus varying as a linear function of the block number (referred to as *block\_delayed*), which will allow analysing the effect of time on choice revaluation effects. For statistical testing, the following linear mixed-effects model will be fitted:

$$choice\_side \sim condition\_side * value\_diff\_side + condition\_side : (reward + reward\_category + reward\_level + delayed\_probe\_phase + block\_delayed) + RE$$

The random effects term *RE* comprises random effects for *subject* (random intercept and random slope for *condition\_side* and *condition\_side : block\_delayed*) and *block* (random intercept; *block* is nested to *subject*). The critical predictor *block\_delayed* codes the distance of the first probe phase

from the delayed probe phase in block units (i.e. *block\_delayed* is 0 for the immediately preceding block, 1 for the block before that block and 2 for the earliest considered block). The binary predictor *delayed\_probe\_phase* indicates the first (*delayed\_probe\_phase* = 0) or the second (*delayed\_probe\_phase* = 1) delayed probe phase. The first main outcome of this analysis will be the strength of the predictor *condition\_side*, thereby testing whether choice revaluation effects are detectable at all in the delayed probe phase. The second main outcome will be the strength of the interaction *condition\_side: block\_delayed*, which quantifies the temporal change of possible revaluation effects.

#### *Self-reinforcement of instrumental choices in the absence of external feedback*

The confidence learning hypothesis posits that confidence in choices presents a learning signal that can further reinforce these choices. An additional test of this hypothesis – although not the focus of this paper – will be to assess whether choices are reinforced during the revaluation phase. In particular, this entails two predictions. The first prediction is that observers become more confident in their choices, even in the absence of feedback. To test this prediction, it will be made use of the fact that each of the three choice pairs in the revaluation phase appear overall three times. A predictor *nchoices* will be introduced, such that for each trial of the revaluation phase, *nchoices* codes how often the current choice has occurred for this particular choice pair (including the current one). The following model will be fitted to the trials of the revaluation phase:

$$confidence \sim nchoices * (value\_chosen + value\_diff\_chosen + reward\_diff\_chosen) + block + reward\_sum + reward\_level + decision\_noise + RE$$

The main outcome will be the strength of the main effect *nchoices*, which quantifies the degree of self-reinforcement across repeated choices of a CS.

A second prediction associated with the confidence learning hypothesis is that self-reinforcement increases the consistency of choices in the absence of feedback. As each choice pair appears three times in the revaluation phase, the average proportion of consistent choices between the 1st and 2nd occurrence of a choice pair will be compared with the analogous proportion between the 2nd and 3rd occurrence. The main outcome is the difference in proportions and the result of an associated two-sample proportions z-test.

#### **Data availability**

All data and materials associated with the experiment (e.g. images of the CS) will be made publicly available at [https://github.com/konfidenz/choice\\_revaluation](https://github.com/konfidenz/choice_revaluation) upon acceptance for publication of the Stage 2 manuscript.

#### **Code availability**

All code used for this study (online experiment, analysis scripts) will be shared on acceptance of the Stage 2 manuscript at [https://github.com/konfidenz/choice\\_revaluation](https://github.com/konfidenz/choice_revaluation).



## References

- Brehm JW (1956) Postdecision changes in the desirability of alternatives. *J Abnorm Soc Psychol* 52:384–389.
- Chen MK (2008) Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Preferences? Cowles Found Discuss Pap No 1669.
- Chen MK, Risen JL (2010) How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm. *J Pers Soc Psychol* 99:573–594.
- Clifford S, Jerit J (2016) Cheating on Political Knowledge Questions in Online Surveys: An Assessment of the Problem and Solutions. *Public Opin Q* 80:858–887.
- Clos M, Schwarze U, Gluth S, Bunzeck N, Sommer T (2015) Goal- and retrieval-dependent activity in the striatum during memory recognition. *Neuropsychologia* 72:1–11.
- Coppin G, Delplanque S, Cayeux I, Porcherot C, Sander D (2010) I'm no longer torn after choice: How explicit choices implicitly shape preferences of odors. *Psychol Sci* 21:489–493.
- Coppin G, Delplanque S, Porcherot C, Cayeux I, Sander D (2012) When Flexibility Is Stable: Implicit Long-Term Shaping of Olfactory Preferences Martinez LM, ed. *PLoS ONE* 7:e37857.
- Daniel R, Pollmann S (2010) Comparing the neural basis of monetary reward and cognitive feedback during information-integration category learning. *J Neurosci* 30:47–55.
- Daniel R, Pollmann S (2012) Striatal activations signal prediction errors on confidence in the absence of external feedback. *NeuroImage* 59:3457–3467.
- Festinger L (1957) *A Theory of Cognitive Dissonance*. Stanford, California: Stanford University Press.
- Fujiwara J, Usui N, Park SQ, Williams T, Iijima T, Taira M, Tsutsui K-I, Tobler PN (2013) Value of freedom to choose encoded by the human brain. *J Neurophysiol* 110:1915–1929.
- Gerard HB, White GL (1983) Post-Decisional Reevaluation of Choice Alternatives. *Pers Soc Psychol Bull* 9:365–369.
- Gershman SJ (2016) Empirical priors for reinforcement learning models. *J Math Psychol*:6.
- Greenwald HJ (1969) Dissonance and relative versus absolute attractiveness of decision alternatives. *J Pers Soc Psychol* 11:328–333.
- Guggenmos M, Wilbertz G, Hebart MN, Sterzer P (2016) Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *eLife* 5:1–19.
- Hamari J, Koivisto J, Sarsa H (2014) Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. In: 2014 47th Hawaii International Conference on System Sciences, pp 3025–3034. Waikoloa, HI: IEEE. Available at: <http://ieeexplore.ieee.org/document/6758978/> [Accessed November 14, 2020].
- Hebart M, Schriever Y, Donner TH, Haynes J-D (2016) The Relationship between Perceptual Decision Variables and Confidence in the Human Brain. *Cereb Cortex* 26:118–130.
- Hoffman MD, Gelman A (2014) The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *J Mach Learn Res* 15:1593–1623.
- Izuma K, Akula S, Murayama K, Wu D-A, Iacoboni M, Adolphs R (2015) A Causal Role for Posterior Medial Frontal Cortex in Choice-Induced Preference Change. *J Neurosci* 35:3598–3606.
- Izuma K, Matsumoto M, Murayama K, Samejima K, Sadato N, Matsumoto K (2010) Neural correlates of cognitive dissonance and choice-induced preference change. *Proc Natl Acad Sci U S A* 107:22014–22019.
- Johansson P, Hall L, Tärning B, Sikström S, Chater N (2014) Choice Blindness and Preference Change: You Will Like This Paper Better If You (Believe You) Chose to Read It!: Choice Blindness and Preference Change. *J Behav Decis Mak* 27:281–289.
- Kaufman EL, Lord MW, Reese TW, Volkman J (1949) The Discrimination of Visual Number. *Am J Psychol* 62:498.
- Koster R, Duzel E, Dolan RJ (2015) Action and valence modulate choice and choice-induced preference change. *PLoS ONE* 10:1–10.

- Lange K, Kühn S, Filevich E (2015) "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies Margulies D, ed. PLOS ONE 10:e0130834.
- Leotti LA, Delgado MR (2011) The inherent reward of choice. *Psychol Sci* 22:1310–1318.
- Leotti LA, Delgado MR (2014) The Value of Exercising Control Over Monetary Gains and Losses. *Psychol Sci* 25:596–604.
- Leotti LA, Iyengar SS, Ochsner KN (2010) Born to choose: the origins and value of the need for control. *Trends Cogn Sci* 14:457–463.
- Luettgau L, Tempelmann C, Kaiser LF, Jocham G (2020) Decisions bias future choices by modifying hippocampal associative memories. *Nat Commun* 11:3318.
- Luo J, Yu R (2017) The Spreading of Alternatives: Is it the Perceived Choice or Actual Choice that Changes our Preference?: Perceived Choice and Actual Choice in our Preference. *J Behav Decis Mak* 30:484–491.
- Magezi DA (2015) Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui). *Front Psychol* 6:2.
- Nakamura K, Kawabata H (2013) I Choose, Therefore I Like: Preference for Faces Induced by Arbitrary Choice Pessiglione M, ed. PLoS ONE 8:e72071.
- Rescorla RA, Wagner AR (1972) A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. In: *Classical conditioning II: current research and theory* (Black AH, Prokasy WF, eds), pp 64–99. New York: Appleton-Century-Crofts.
- Salvatier J, Wiecki TV, Fonnesbeck C (2016) Probabilistic programming in Python using PyMC3. *PeerJ Comput Sci* 2:e55.
- Schwarze U, Bingel U, Badre D, Sommer T (2013) Ventral Striatal Activity Correlates with Memory Confidence for Old- and New-Responses in a Difficult Recognition Test. PLoS ONE 8:e54324.
- Sharot T, Fleming SM, Yu X, Koster R, Dolan RJ (2012) Is Choice-Induced Preference Change Long Lasting? *Psychol Sci* 23:1123–1129.
- Sharot T, Martino BD, Dolan RJ (2009) How choice reveals and shapes expected hedonic outcome. *J Neurosci* 29:3760–3765.
- Sharot T, Velasquez CM, Dolan RJ (2010) Do decisions shape preference? Evidence from blind choice. *Psychol Sci* 21:1231–1235.
- Shultz TR, Léveillé E, Lepper MR (1999) Free Choice and Cognitive Dissonance Revisited: Choosing "Lesser Evils" Versus "Greater Goods." *Pers Soc Psychol Bull* 25:40–48.
- Skipper S, Perktold J (2010) statsmodels: Econometric and statistical modeling with python. In: 9th Python in Science Conference.
- Steele CM, Spencer SJ, Lynch M (1993) Self-image resilience and dissonance: the role of affirmational resources. *J Pers Soc Psychol* 64:885–896.
- van Geen C, Gerraty RT (2020) Hierarchical Bayesian Models of Reinforcement Learning: Introduction and comparison to alternative methods. *BioRxiv* Available at: <http://biorxiv.org/lookup/doi/10.1101/2020.10.19.345512> [Accessed January 18, 2021].
- Voigt K, Murawski C, Speer S, Bode S (2020) Effective brain connectivity at rest is associated with choice-induced preference formation. *Hum Brain Mapp* 41:3077–3088.
- Watanabe S (2010) Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *J Mach Learn Res* 11:3571–3594.
- White G (1981) Postdecision evaluation of choice alternatives as a function of valence of alternatives, choice, and expected delay of choice consequences\*1. *J Res Personal* 15:371–382.

## Acknowledgements

This study was supported by the grant GU 1845/1-1 of the German Research Foundation (DFG). The funders have/had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions

M.G. conceived the study. I.S., P.S. and M.G. designed the study. I.S., P.S. and M.G. drafted the stage 1 manuscript.

## Competing interests

The authors declare no competing interests.

## Design Table

Question	Hypothesis	Sampling plan	Analysis plan	Interpretation given to different outcomes
Does choice revaluation generalize to instrumental conditioning?	Choice revaluation leads to an increase of subjective value ratings for CS of set 2 (higher choice probability in the revaluation phase) relative to CS of set 1 (lower choice probability in the revaluation phase)	According to the power calculation a sample size of N=168 is required to achieve an a priori power of 0.95 for the minimal effect of interest of the hypothesis that is expected to have the smallest effect size (see Section Participants and online experiment). We expect a dropout of approximately 10% and will resample until the desired sample size is achieved. All design contrasts will be within-subject and will be tested on the same population of participants.	<u>Confirmatory planned analysis 1</u> A linear mixed-effects model will be used to test for a significant main effect of <i>condition</i> (set 1/set 2) on subjective value ratings consistent with the hypothesis that subjective values increase for CS of set 2 relative to CS of set 1. See section “Confirmatory planned analysis 1” for a full specification of the model.	A significant positive coefficient of the main effect <i>condition</i> is interpreted as evidence in favor of the hypothesis. Otherwise, it will be concluded that there is not sufficient evidence for the hypothesis.
	Choice revaluation leads to preferential instrumental choices for CS of set 2 (higher choice probability in the revaluation phase) relative to CS of set 1 (lower choice probability in the revaluation phase)		<u>Confirmatory planned analysis 2</u> A logistic mixed model will be used to test for a significant main effect of <i>condition_side</i> (set 1/set 2) on instrumental choices consistent with the hypothesis that observers are more likely to choose CS of set 2 relative to CS of set 1. See section “Confirmatory planned analysis 2”, for a full specification of the model.	A significant positive coefficient of the main effect <i>condition_side</i> is interpreted as evidence in favor of the hypothesis. Otherwise, it will be concluded that there is not sufficient evidence for the hypothesis.
Can conventional reinforcement learning models be extended with a choice revaluation term that explains the dynamics of choice revaluation?	The best-fitting reinforcement learning model with a choice revaluation term has higher model evidence than a null model without choice revaluation.		<u>Confirmatory planned analysis 3</u> Bayesian model comparison using the Watanabe-Akaike information criterion (WAIC). A model is declared superior to another if the corresponding 95% confidence intervals of WAIC do not overlap.	If no choice revaluation model outperforms the null model, it will be concluded that choice revaluation – if it exists – is not meaningfully captured by any of the tested models.
How does choice confidence affect choice revaluation?	Model-based contrast of three hypotheses: (1) confidence does not modulate choice revaluation ( <i>mere-choice hypothesis</i> ), (2)			If one of the three hypothesis-driven models outperforms both others, the associated hypothesis will be accepted. If two of the models outperform a third model, but are themselves indistinguishable, the




	confidence negatively modulates choice revaluation ( <i>cognitive-dissonance hypothesis</i> ), (3) confidence positively modulates choices revaluation ( <i>confidence-learning hypothesis</i> )		hypothesis associated with the third model will be rejected, but neither specific hypothesis of the winning models will be accepted. If no model is superior, it will be concluded that there is not sufficient evidence for either hypothesis.
--	---	--	---

## Supplementary information


### Instructions

#### The story so far...

After days of following a mysterious map, you and your crew finally found the ancient temple that is said to hold a great lost treasure. The legend holds that only the bravest and smartest explorers will be able to collect the whole treasure hidden in the old ruin.

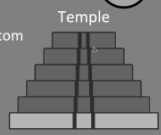


Now is your time to prove yourself! Are you ready? Before you start your adventure, we'll brief you about some basics.




#### Six levels to fame and glory


The temple has six levels and you'll work your way from bottom to top.



In every level you will repeatedly encounter two switches on a wall, each of them with a symbol on top. An example is depicted on the left.




Each switch is associated with a coin vault, which is invisibly located behind the wall. When you press a switch, the coins are released from the vault and you collect them.



Note: pressing one switch blocks the other switch, hence you always have to make a choice between the two.

#### Learning a forgotten language

Your goal is to collect as many coins as possible. Importantly, the amount of coins released when pressing a switch may depend on its symbol. In other words, you have to learn about the symbols.



Unfortunately, the symbol language spoken by the builders of the temple is long lost. You have to learn yourself how valuable the different symbols are.


In each level of the temple, a new set of symbols appears, hence your learning starts anew.

Hint: there are no general patterns such as "all red symbols are more valuable" or "all animal symbols are less valuable". Instead, you have to learn about each symbol by its own.

#### Your crew will have questions...

Each time you make a choice, your companions want to know how sure you are that you chose the vault that holds more coins.

Try to be as truthful as possible. The coins you obtain do not depend on the level of confidence you indicate.



You provide your confidence on a scale from 0 to 5. A value of 0 means that you have no idea which one is the better switch, i.e. you are guessing. A value of 5 means that you're 100% certain you've made the right choice.

Especially in the beginning you will often have to guess, but usually you get more confident along the way. Try to use the full confidence scale to also indicate intermediate levels of confidence.

On the next screen, you can try it out.

## Rating time

5

Every now and then, your companions inquire how valuable the symbols feel to you at this point of the journey.

You provide your rating on a continuous scale, as displayed on the right. The more valuable a symbol feels to you, the more you'd move the slider to the right. Use the rating intuitively.



You don't have to practise this one. You will use the mouse to move the slider and then left-click to select.



## The King's chamber

6

At the end of each level, you will find the "King's Chamber".

The chamber is special in that you do not see the coins directly after each choice. Instead, at the end of the chamber, an additional switch appears to release all coins at once. But note: the temple is generous and triples up the number of coins you earned in the chamber.

The symbols remain the same in the chamber and also they keep their meaning. However, the switches have golden a shade, indicating that they're extra valuable due to the temple's generosity.

