UNIVERSITEIT VAN AMSTERDAM

The Use of a Probabilistic Reinforcement Learning Task and Eye Measurements to Study the Somatic Marker Hypothesis

June 27, 2025

Student:
Siyi Zeng
UvAnetID 14880067

Supervisor:
Ivan Ivanchei
Assessor:
Filip van Opstal
Examiner:
Simon van Gaal

Abstract

The somatic marker hypothesis proposes that physiological signals guide decision-making under uncertainty by biasing choices toward advantageous outcomes. However, traditional paradigms such as the Iowa Gambling Task, along with physiological measures like skin conductance, have methodological limitations that obscure the precise mechanisms underlying somatic marker function. This study addresses these limitations by combining probabilistic reinforcement learning, pupillometry, and computational modeling to investigate the relationship between pre-choice physiological responses and value-based decision-making. Forty-four healthy participants performed a probabilistic reinforcement learning task with sequential stimulus presentation while pupil responses were recorded. Bayesian hierarchical Q-learning models provided trial-by-trial estimates of value beliefs. Results revealed three key findings: First, stimulus-related pupil responses encoded value difference information between options rather than reward magnitude, with larger pupil dilation for smaller value differences between options. Second, choice-related analyses demonstrated systematic correlations between pupil responses and computational Q-values throughout the decision period. Third, and most critically, linear mixed-effects analyses revealed enhanced pupil sensitivity to value differences on accurate versus inaccurate trials across both pre- and post-choice periods. This differential value-autonomic coupling provides evidence for somatic marker function while revealing that these signals operate through temporally dynamic mechanisms rather than static warnings. These findings support a refined model of the somatic marker hypothesis with important implications for understanding decision-making processes through pupillometric measures.

Keywords: Somatic marker hypothesis; Pupillometry; Reinforcement learning; Decision-making; Computational modeling; Autonomic nervous system; Uncertainty

Introduction

Daily experience continuously requires us to make decisions under conditions of uncertainty, from trivial choices like selecting a lunch spot to more consequential ones such as choosing a financial investment. In such contexts, humans frequently rely on instinctive, affect-laden judgements—commonly referred to as "gut feelings". These decisions, while seemingly irrational, may be guided by physiological and emotional signals that reflect prior experiences and anticipated outcomes. This view is encapsulated in the somatic marker hypothesis (SMH), originally proposed by Antonio Damasio, which posits that physiological states linked to emotional experiences—somatic markers—can bias decision-making towards advantageous outcomes (Damasio, 1996).

SMH was initially validated using the Iowa Gambling Task (IGT), a paradigm designed to simulate real-life decision-making by incorporating uncertainty, delayed reward, and punishment (Bechara et al., 2005; Damasio, 1996). Research using the IGT demonstrated that healthy individuals, even without explicit knowledge of the task structure, tend to gradually favor advantageous options (Becker et al., 2019; Damasio, 1996). This behavior was accompanied by anticipatory skin conductance responses (SCRs) that appeared before making a choice, particularly when considering disadvantageous options (Bechara et al., 2005; Damasio, 1996). Patients with ventromedial prefrontal cortex (VMPFC) damage, in contrast, lacked these anticipatory SCRs and consistently made poor decisions, suggesting that somatic markers may operate beneath conscious awareness to guide behavior (Bechara et al., 1997; Damasio, 1996).

Despite its influence, the IGT and the use of SCRs have been criticized for both methodological and conceptual reasons (Chiu et al., 2008; Lin et al., 2007; Maia & McClelland, 2004; Tomb et al., 2002). The IGT is a complex task that may require the involvement of several latent cognitive processes. The task's self-paced design hinders fine-grained analysis of dynamic

physiological responses, while SCRs themselves lack temporal and valence specificity (Guillaume et al., 2009). Moreover, alternative explanations have emerged regarding the mechanisms guiding choice behavior during the IGT, suggesting that immediate gain-loss frequency may weigh more heavily than expectations of long-term outcomes (Chiu et al., 2008; Lin et al., 2007; Steingroever et al., 2013). These concerns have prompted calls for new approaches that allow more precise tracking of decision dynamics and physiological signals.

Recent methodological advances suggest a promising direction: reinforcement learning (RL) paradigms provide better experimental control (Van Slooten et al., 2018) as a more analytically tractable alternative to the IGT, while pupillometry offers superior temporal resolution compared to SCRs. Specifically, RL tasks can offer clearer separation between stimulus presentation, choice, and feedback phases, facilitating better temporal alignment of physiological and behavioral data. Within this framework, a new generation of research has begun investigating pupil dilation in decision-making (Bierman et al., 2004; Simonovic et al., 2017; Van Slooten et al., 2018; Xu & Huang, 2020). Pupil size, governed by the autonomic nervous system, is a sensitive index of arousal, attention, and cognitive effort (Mckinnon et al., 2020; Steinhauer et al., 2004). It has been shown to correlate with reward prediction errors, value sensitivity, and surprise in decision-making tasks (Cavanagh et al., 2014; de Gee et al., 2014; Hess & Polt, 1964; Lavín et al., 2014; Lempert et al., 2015; Van Slooten et al., 2018; Wang et al., 2018).

Pupil responses have already been linked to reinforcement learning processes (Guath et al., 2023; Kozunova et al., 2022; Van Slooten et al., 2018). Studies show that pupil dilation varies as a function of value beliefs and exploration-exploitation strategies. For example, Van Slooten et al., (2018) conducted probabilistic reinforcement learning tasks in healthy participants and used computational modelling to discover that the pre-choice pupil dilation was a function of trial-by-

trial fluctuations in value beliefs about the upcoming choice and predicted an individual's tendency to exploit or explore. However, the specific hypothesis proposed by SMH—that somatic signals during the evaluation phase of a decision can predict the chosen option and its rewarding nature—has never been directly tested using pupil dilation.

Here, we aim to fill this gap by testing whether pupil dilation shows differential value sensitivity between correct and incorrect decisions in a simple probabilistic reinforcement learning task. We measured pupil size while forty-four participants performed a probabilistic RL task, with separate viewing stages for each stimulus within trials. (Fig.1 and Methods). Our task design enables us to isolate the physiological response to each option, allowing a within-trial comparison of pupil dilation before a decision is formed. Such a design provides a critical advantage over previous paradigms by directly linking stimulus-level physiological responses to decision quality and optimizing temporal alignment between somatic activity and decision events.

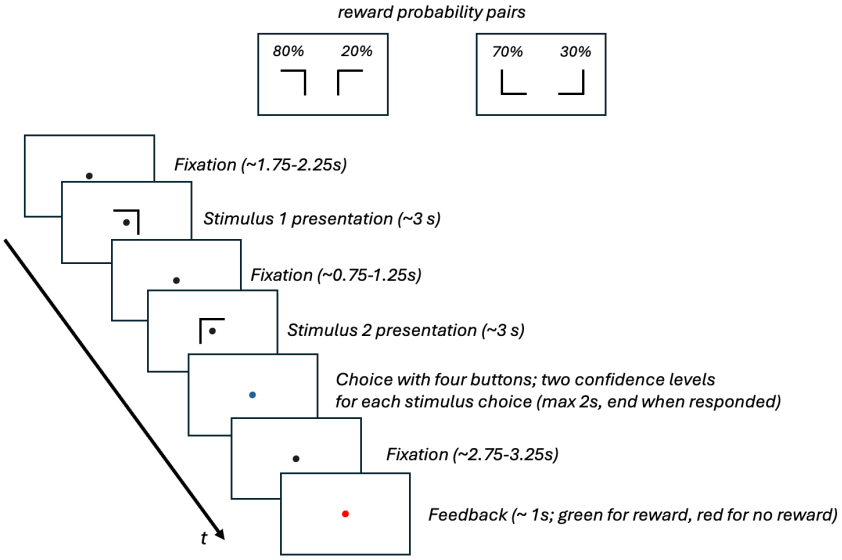


Figure 1. Probabilistic reinforcement learning task design. During each trial, two stimuli from predefined pairs were presented sequentially in randomized order, followed by choice and feedback phases. Participants learned to select the more rewarding option within each pair through probabilistic feedback (green dot = reward, red dot = no reward). The reward probability pairs (80:20 and 70:30) and stimulus-probability assignments were randomized across participants.

We fitted an adapted Bayesian version of the Q-learning RL algorithm to participants' choices to describe value-based decisions and outcome evaluations, following Van Slooten et al., (2018). This computational approach allowed us to evaluate whether pre-choice somatic responses, as indexed by pupil size during option viewing and prior to decision, correlate with the values of the viewed stimulus and upcoming choice, which are potentially associated with which option participants will choose and whether that choice will be optimal.

This study focused on two primary questions: (1) How do stimulus-related pupil responses correlate with the rewarding nature of the stimulus, including its reward probability and its learned value beliefs? (2) How do choice-related pupil responses correlate with the correctness of decisions, as well as the learned beliefs about the chosen and unchosen options? By leveraging pupillometry and reinforcement learning with computational modelling, we aim to provide fresh insights into the role of pre-choice bodily signals in guiding behavior under uncertainty.

Materials and Methods

Ethics Statement

The Ethics Committee of the Universiteit van Amsterdam approved this study.

Participants

Fifty-nine healthy participants with normal or corrected-to-normal vision completed the experiment (age range = 18-35 years). They were compensated €15 or 1.5 credits for 1.5 hours of participation. Written informed consent was obtained from all participants. Fifteen participants were excluded from analyses due to the overall accuracy being lower than 0.6 (N=14) and missing pupil data (N=1), resulting in a final sample of 44 participants.

Task and Procedure

Participants were seated in a dimly lit, silent room with their head positioned in a chin rest approximately 60 centimeters from the computer screen. After providing informed consent and receiving brief task instructions, participants completed a 10-trial practice session followed by 4 blocks of 60 trials each (240 trials total, 120 presentations of each stimulus pair), with brief breaks between blocks. After each block, participants reported their believed reward probability for each stimulus using a 0.0-1.0 sliding scale, and their earned points were displayed (10 points for optimal choices, 0 for suboptimal choices).

Stimuli and Trial Structure

As shown in Fig.1, the task used four different simple images assigned to different reward probabilities. Two stimulus pairs were created: one with 80:20 reward probabilities and another

with 70:30 probabilities. Image-probability assignments and presentation order within each pair were randomized across participants.

Each trial began with participants fixating on a central black dot. After 200 ms (SD = 25ms), the first stimulus appeared beside the fixation dot for 300ms, followed by a 100ms fixation interval (SD = 25ms) before the second stimulus was presented for 300ms. The choice interval (maximum 200ms) began when the central dot turned blue, ending when participants pressed a key to indicate their choice and confidence level (D/F for first stimulus with/without confidence; J/K for second stimulus with/without confidence). Non-responses triggered a "no response, please pay attention" message. After the choice interval, the fixation dot remained black for 300ms (SD = 25ms) before turning green (reward) or red (no reward) for 100ms as feedback.

Computational Model

Choices were fit with a reinforcement learning ("Q-learning") model adapted from Van Slooten et al., (2018). The model estimates expected values (Q-values) for each option based on individual sequences of choices and outcomes. All Q-values were initialized to 0.5 and maintained cumulatively across all experimental blocks to capture learning transfer.

The Q-value for option *i* on the trial *t+1* is updated according to:

$$Q_i(t+1) = Q_i(t) + \begin{cases} \alpha_{Gain}[r_i(t) - Q_i(t)], if\ r = 1 \\ \alpha_{Loss}[r_i(t) - Q_i(t)], if\ r = 0 \end{cases}$$

Where parameters $0 \leq \alpha_{Gain}, \alpha_{Loss} \leq 1$, respectively represent positive and negative learning rates. Choice probability (e.g., selecting A over B) was described by a SoftMax rule:

$$P_A(t) = \frac{\exp\left(\beta \cdot Q_A(t)\right)}{\exp\left(\beta \cdot Q_B(t)\right) + \exp\left(\beta \cdot Q_A(t)\right)}$$

Where the explore-exploit parameter, $0 \leq \beta \leq 10$, described the sensitivity to option value differences, with larger β values indicating greater sensitivity to value differences. The β parameter range was constrained to 0-10 for improved numerical stability compared to wider ranges.

Bayesian Hierarchical Modelling Procedure

The Q-learning model was fitted using Bayesian hierarchical procedures, where individual parameter estimates were drawn from group-level parameter distributions that constrained the range of possible individual parameter estimates (Van Slooten et al., 2018). Our model was implemented with optimizations for the 44-participant dataset. Variables $r_i(t-1)$ (reward outcome for participant *i* on trial *t-1,* and $ch_i(t)$ (choice of participant *i* on trial *t*) were obtained from the behavioral data. Per-participant parameter estimates $\alpha_{Gain}, \alpha_{Loss}, \beta$ were modelled using a non-centered parameterization with probit transformation. The hierarchical priors were adjusted for our larger sample size: group-level means $\mu \sim \mathrm{N}(0, 0.8)$ and group-level standard deviations $\sigma \sim \mathrm{N}(0, 0.5 - 0.7)$, allowing for enhanced individual variation modelling while maintaining stability. The Bayesian hierarchical model was implemented using STAN, a probabilistic programming language for Bayesian statistical inference (Carpenter et al., 2017), and fit to all trials. The model was compiled and executed using CmdStanPy, the Python interface to CmdStan, which provides access to Stan's Hamiltonian Monte Carlo (HMC) sampling with the No-U-Turn Sampler (NUTS) algorithm to efficiently explore the posterior distribution of model parameters. 4 chains, with 2000 sampling and 2000 warmup iterations per chain, were generated to ensure convergence, which was evaluated by the Rhat statistic.

Our adapted Q-learning model underwent comprehensive validation to ensure parameter identifiability and predictive validity. Parameter recovery testing confirmed recovery of known

parameters from simulated data (mean correlations: β r = 0.804, α_gain r = 0.752, α_loss r = 0.768). Posterior predictive validation demonstrated strong correspondence between empirical and simulated learning curves at both group (r = 0.997) and individual levels (mean individual correlation r = 0.635, 32/44 subjects with r > 0.6). Model fit assessment yielded WAIC = 9182.6 with 0.0% divergent transitions, confirming reliable parameter estimation and good model quality for scientific inference.

Pupillometry: Preprocessing

Pupil diameter was recorded at 1000 Hz using an EyeLink 1000 Tower Mount (SR Research). The eye tracker was calibrated before each experimental block. Raw pupil data were pre-processed using a multi-stage pipeline implemented in Python with MNE-Python for eye-tracking data handling.

Blink Detection and Interpolation. Blinks and saccades were detected using standard EyeLink software with default settings. Periods of data loss during blinks were interpolated using MNE's blink interpolation function with a buffer window of 100 ms before and after each detected blink. This linear interpolation approach replaced missing data points to maintain signal continuity across the time series.

Temporal Filtering. The interpolated pupil signal was temporally filtered using third-order Butterworth filters. A low-pass filter with a cutoff frequency of 10 Hz was applied to remove high-frequency noise, followed by a high-pass filter with a cutoff frequency of 0.01 Hz to remove slow drifts and create a bandpass-filtered signal.

Signal Processing. The cleaned pupil signal was converted to a percent signal change relative to the median pupil size within each experimental block. Data were down-sampled to 100

Hz for analysis and epoched around key experimental events (stimulus presentation, choice, and feedback) with appropriate baseline correction windows.

Statistical Analysis: Stimulus-Related Effects

To examine the stimulus-related effects on pupil responses, we conducted a series of analyses comparing pupil dynamics across different reward and value belief conditions. All analyses employed participant-level aggregation to avoid pseudo-replication, with individual trial data first averaged within each participant and condition before statistical testing. Such analyses have been performed on both stimulus-1 and stimulus-2 pupil responses.

Reward-Based Pupil Response Analysis

High vs. Low Reward. We first examined pupil responses to stimuli associated with different reward probabilities (rp). Trials were categorized into high reward (rp = 0.7, 0.8) and low reward (rp = 0.2, 0.3) conditions. Pupil responses were aggregated at the participant level for each condition, then subjected to nonparametric cluster-based permutation testing. For each time point, paired t-tests were performed comparing high vs. low reward conditions across participants. Clusters were identified as contiguous periods where $|t| > 2.0$ ($p < 0.05$). The observed cluster statistics (sum of t-values within each cluster) were compared against a null distribution generated by 1000 permutations of condition labels within participants.

Value Belief Analysis (Q_stim). We investigated whether pupil responses reflected learned value beliefs about the stimulus. Q_stim values (value beliefs of the stimulus) were extracted from computation model outputs and z-scored within each participant to control for individual value differences in value scaling. Two complementary approaches were used: (1) linear mixed-effects modelling with pupil response as the dependent variable, standardized Q_stim as

the fixed effect, and random intercepts by participant; (2) Within-subject median split analysis, where Q_stim values were divided into high and low conditions based on each participant's median, followed by cluster-based permutation testing as described above.

Value Difference Analysis

Large vs. Small Value Difference. To test pupil's sensitivity to the value difference magnitude between visual stimuli, we compared pupil responses between trials with large value difference (stimulus pair 80:20) and small value difference (stimulus pair 70:30). This analysis employed the same cluster-based permutation approach, testing whether pupils are sensitive to the value difference within the stimulus pair.

Value Difference Belief Analysis (Q_diff). We examined whether pupil responses encoded the magnitude of value difference between choice options. Q_diff values, |Q_chosen - Q_notchosen|, were extracted from computational models and z-scored within participants. As with Q_stim, we employed both linear mixed-effects modelling and within-subject median split permutation testing.

Statistical Analysis: Choice-Related Effects

Following our stimulus-related analyses, we examined how pupil responses tracked value-based decision-making processes during choice formation and evaluation. These analyses employed the same cluster-based permutation testing framework established for our stimulus analyses and were adapted based on Van Slooten et al., (2018).

Q-value Correlation Analysis

We tested whether pupil responses correlated with trial-wise Q-values derived from our computational model. For each participant and time point, Pearson correlations were calculated

between pupil diameter and three Q-value components: (1) Q_chosen: value beliefs about the selected option; (2) Q_unchosen: value beliefs about the non-selected option; (3) Q_diff: absolute difference between chosen and unchosen values.

Nonparametric cluster-based permutation t-tests were used to test for significant correlations between pupil responses and Q-values while correcting for multiple comparisons over time. Briefly, for each time point, t-tests were performed on correlation coefficients across participants to test whether correlations differed significantly from zero. The observed cluster statistics were then compared to a null distribution generated by within-subject permutation: for each of the 1000 permutations, Q-values were randomly shuffled across trials within each participant, preserving individual differences while breaking trial-specific pupil-value relationships. The proportion of permuted maximum cluster statistics exceeding the observed cluster statistic determined the corrected P-value.

Choice Accuracy Analysis

To test the somatic marker hypothesis by examining whether pre-choice pupil responses predict decision correctness, we compared pupil dynamics between accurate (choosing the larger reward probability option) and inaccurate (choosing the lower reward probability option) trials. Within each participant, trials were aggregated by accuracy condition, and mean pupil responses were calculated for each condition. Accuracy labels were randomly permuted within participants for each permutation, with particular focus on pre-choice time windows (-2.0s to 0s) to test the predictive nature of somatic markers.

Q-value Accuracy Interaction Analysis

To test whether pupil sensitivity to Q-values systematically differed between accurate and inaccurate trials during decision formation, we conducted linear mixed-effects analyses examining

Q-value × accuracy interactions. This approach directly explores the somatic marker hypothesis by examining whether stronger value-autonomic coupling correlates with decision quality. For each time point, we fitted LME models with the formula:

$$\text{pupil} = \beta_0 + \beta_1(Q\_value\_z) + \beta_2(accuracy) + \beta_3(Q\_value\_z \times accuracy) + error,$$

Where the interaction term tests whether Q-value sensitivity differs by trial accuracy. Here, the critical parameter was the interaction coefficient β_3, representing the difference in Q-value sensitivity on accurate trials. Significant interactions were identified using cluster-based permutation testing, where accuracy labels were shuffled within participants across 1000 permutations to generate null distributions. For significant clusters, we conducted simple effects analyses by fitting separate models for correct and incorrect trials to quantify the Q-value sensitivity difference between accuracy conditions.

Results

Behavioral and Model Performance

Participants successfully learned the reward contingencies, showing above-chance accuracy that improved across blocks (Fig.2A, 2B). The Q-learning model accurately captured individual choice behavior, with strong correspondence between empirical and simulated learning curves (r = 0.997, Fig.2B). Fitted parameters showed reasonable distributions (Fig.2C), and model-estimated Q-values appropriately differentiated reward probabilities (Fig.2D).

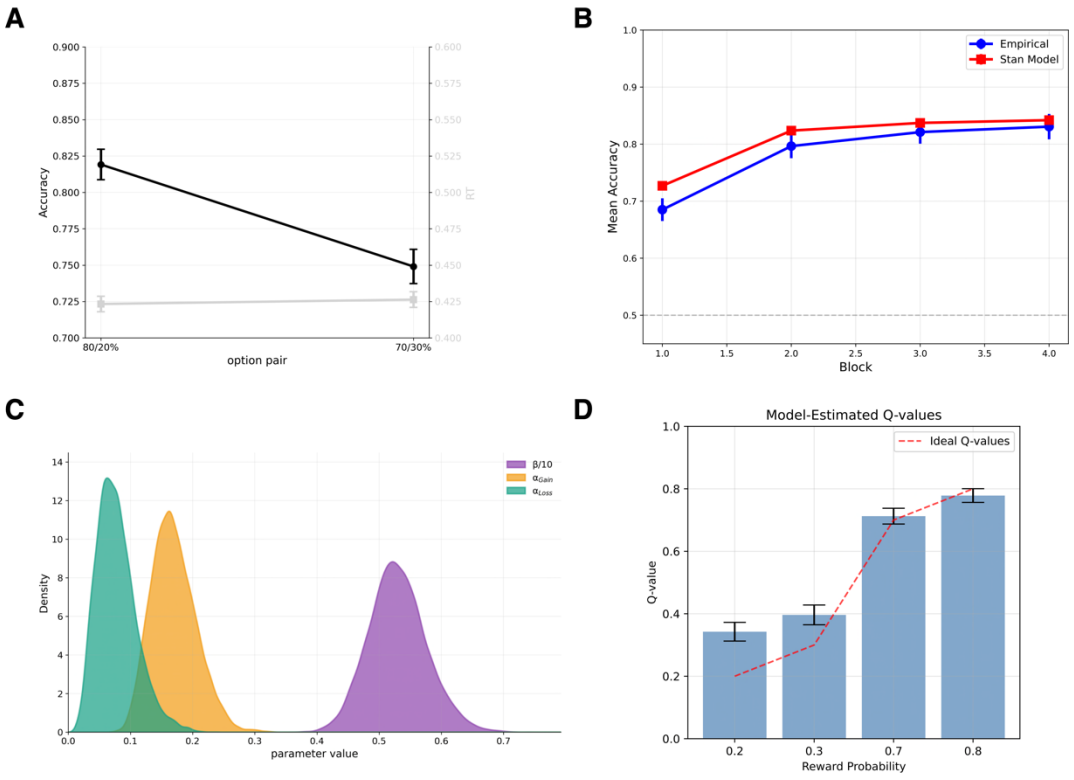


Figure 2. Behavioral performance and computational model validation. A. Mean accuracy and reaction times across participants (N=44) as a function of option pair difficulty. **B.** Empirical (blue) and model-simulated (red) accuracy across experimental blocks, demonstrating strong correspondence between observed and predicted learning curves (r = 0.997). **C.** Group-level posterior distributions for fitted Q-learning parameters: α_{Gain} (positive learning rate), α_{Loss} (negative learning rate), and β (exploration-exploitation parameter). **D.** Model-estimated Q-values for each reward probability condition, showing appropriate value differentiation corresponding to objective reward rates. $\beta/10$ scaling applied for visualization clarity.

Stimulus-related pupil responses reflect value difference rather than reward magnitude

We first examined whether pupil responses during stimulus presentation differentiated between high reward and low reward stimuli (Fig.3). While high reward stimuli showed trends toward larger pupil responses during stimulus 1 presentation, no significant clusters emerged after multiple comparison correction. Value beliefs about individual stimuli (Q_stim1, Q_stim2) also showed no correlations with pupil responses (see S1 Fig).

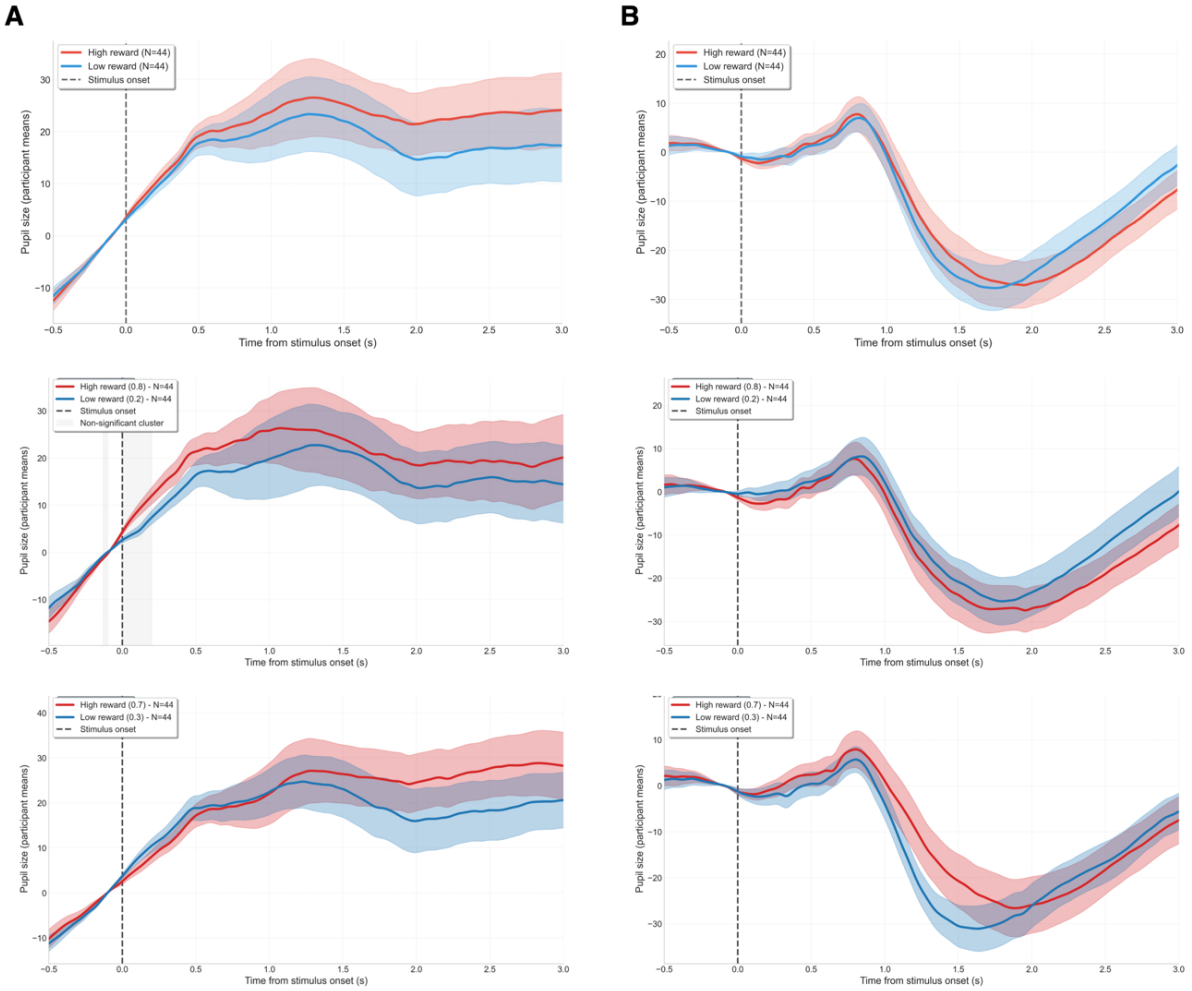


Figure 3. Stimulus-related pupil responses do not differentiate reward magnitude. A. Stimulus-1 pupil responses split by reward probability conditions. Upper panel: High reward (0.7, 0.8) vs low reward (0.2, 0.3) comparison showing non-significant trends toward larger responses for high-reward stimuli. Middle panel: 0.8 vs 0.2 showing a similar non-significant pattern. Lower panel: 0.7 vs 0.3 showing minimal differentiation. **B.** Stimulus-2 pupil responses using identical reward probability groupings, showing weaker and less consistent patterns than Stimulus-1. Shaded areas represent standard error across participants. No significant clusters emerged after multiple comparison correction, indicating limited sensitivity to individual stimulus reward magnitude.

In contrast, stimulus-1-related pupil responses strongly encoded value differences between choice options. Q_diff values (representing learned beliefs about absolute value difference between options) correlated significantly with pupil responses around 2s after stimulus 1 onset (Fig.4A, upper and middle panels). The relationship was predominantly negative, indicating that smaller value differences—representing greater decision difficulty—were associated with larger pupil responses. Cluster-based permutation testing further identified a significant cluster during later stimulus presentation (2.5s post-onset), where low Q_diff trials showed consistently larger pupil responses than high Q_diff trials. These patterns suggest that pupils primarily encode value differences between options rather than individual stimulus reward magnitude during the evaluation phase. This negative relationship aligns with cognitive effort theories, where smaller value differences require greater cognitive effort due to increased choice difficulty (van der Wel & van Steenbergen, 2018). Stimulus-2-related pupil responses showed less consistent results (Fig.4B, upper and middle panels), with a later significant cluster, which we considered related to pre-choice dilation and discussed in the next section. However, the practical value difference, or the condition of the stimulus pair, did not show significant correlation with stimulus-related pupil responses (Fig.4, lower panels).

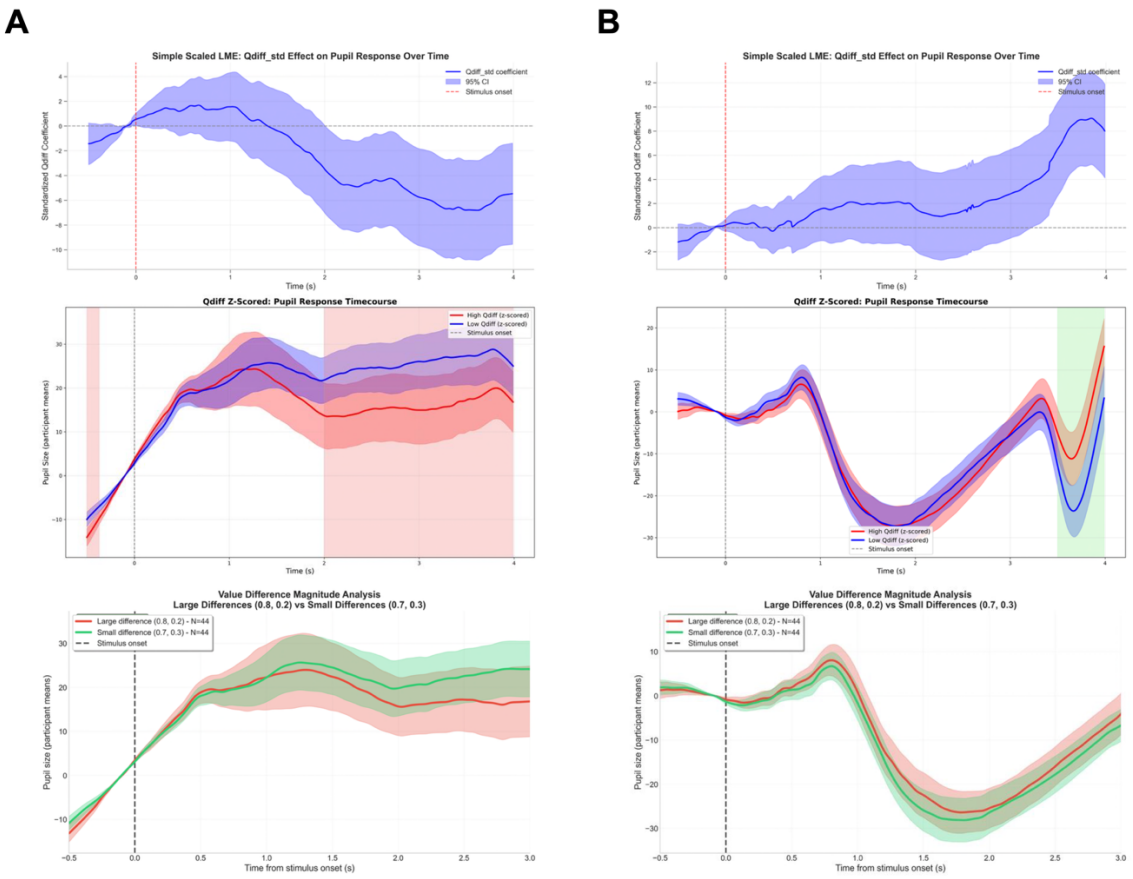


Figure 4. Stimulus-related pupil responses encode the value difference between options. **A.** Q_diff correlation with stimulus-1 pupil responses. Upper panel: Linear mixed-effects model coefficients showing a significant negative relationship (mean coefficient = -2.336), indicating larger pupil responses for smaller value differences or more difficult decisions. Middle panel: Within-subject permutation analysis comparing high vs low Q_diff trials, with a significant cluster at 2.5s post-stimulus onset (red shaded area, $p < 0.05$). Lower panel: Within-subject permutation analysis comparing trials with 80:20 stimulus pair vs. 70:30 stimulus pair, with no significant cluster. **B.** Q_diff correlation with stimulus-2 pupil responses showing a weaker but significant late cluster. No significant difference in stimulus-2-related pupil responses was found between the two stimulus pair conditions.

Choice-related pupil responses reveal enhanced value sensitivity on accurate trials

Choice-related pupil dynamics showed significant correlations with learned Q-values throughout the decision period. Q_chosen (value of selected option) correlated positively with pupil responses from pre-choice through post-choice intervals (Fig.5A), while Q_unchosen (value of non-selected option) showed weaker negative correlations primarily post-choice (Fig.5B). Q_diff (absolute difference between option values) demonstrated consistent positive correlations reflecting decision difficulty sensitivity (Fig.5C). Notably, this positive correlation with Q_diff during the choice

period contrasts with the negative correlation observed during stimulus evaluation. These temporal dynamics align with results from Boersma et al., (1970) and indicate that somatic markers could serve different functions at different stages of decision-making, signalling cognitive effort during evaluation and value differentiation during choice implementation.

Direct comparison between accurate and inaccurate trials revealed small but consistent differences, with correct trials showing larger pupil responses (Fig.5D). However, no significant clusters survived multiple comparison correction, indicating pupil responses do not provide simple binary signals for decision accuracy.

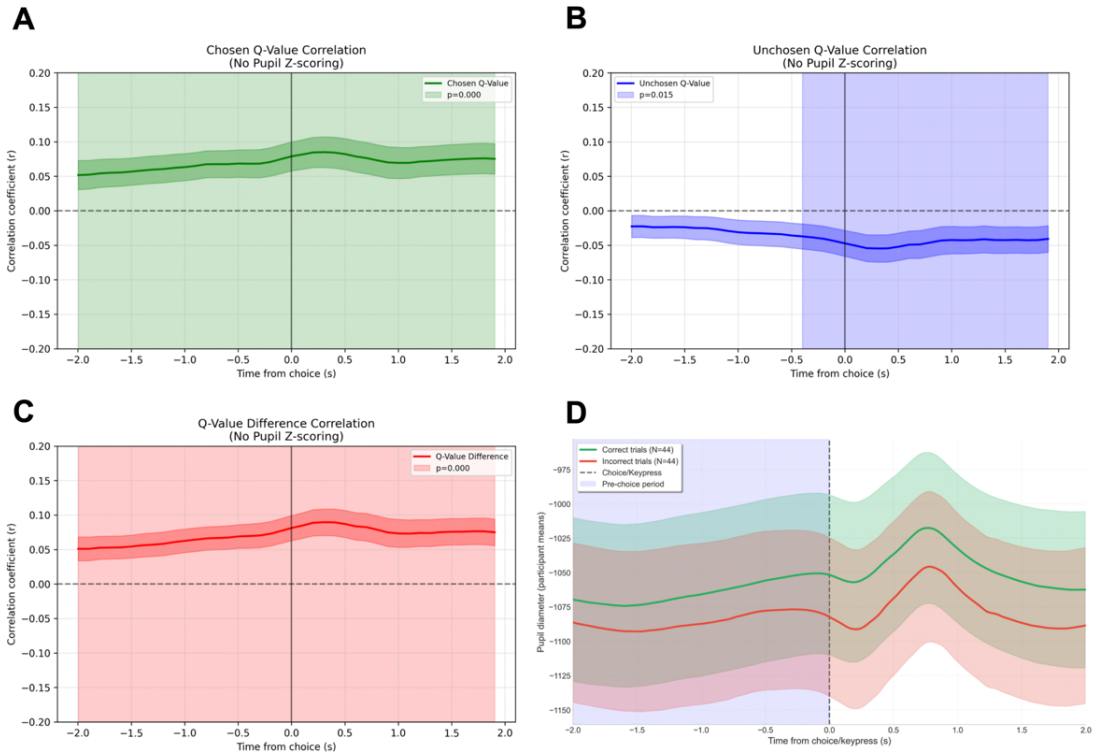


Figure 5. Choice-related pupil dilation revealed value sensitivity, not decision accuracy. A-C. Correlations between pupil diameter and computational Q-values across the choice epoch (-2 to +2s relative to choice). Q_chosen shows sustained positive correlations throughout the decision period (A), Q_unchosen exhibits weaker negative correlations primarily post-choice (B), and Q_diff demonstrates consistent positive correlations reflecting decision difficulty sensitivity (C). Shaded areas indicate significant clusters after cluster-based permutation correction ($p < 0.05$). **D.** Direct comparison of pupil responses between accurate (green) and inaccurate (red) trials showing small but consistent differences with no significant clusters after multiple comparison correction, indicating limited binary predictive power for decision accuracy.

Linear mixed-effects interaction analyses revealed that pupil sensitivity to Q-values systematically differed between accurate and inaccurate trials, providing direct evidence for differential value-autonomic coupling. For Q_chosen, a significant post-choice cluster (0.1-0.9s) showed enhanced value sensitivity on correct trials (Fig.6A), suggesting stronger value-autonomic coupling during the outcome evaluation.

Most critically, Q_diff showed enhanced sensitivity on accurate trials across the entire analysis window (-2.0 to 1.9s), spanning both pre- and post-choice periods (Fig.6B). On correct trials, larger value differences elicited stronger pupil responses, while on incorrect trials, this relationship was attenuated or reversed. This differential sensitivity represents evidence for somatic marker function: appropriate autonomic responses to value difference (Q_diff) systematically differ between correct and incorrect trials, with these differences evident before decisions are made, supporting the hypothesis that physiological signals show differential coupling with value information during adaptive decision-making under uncertainty.

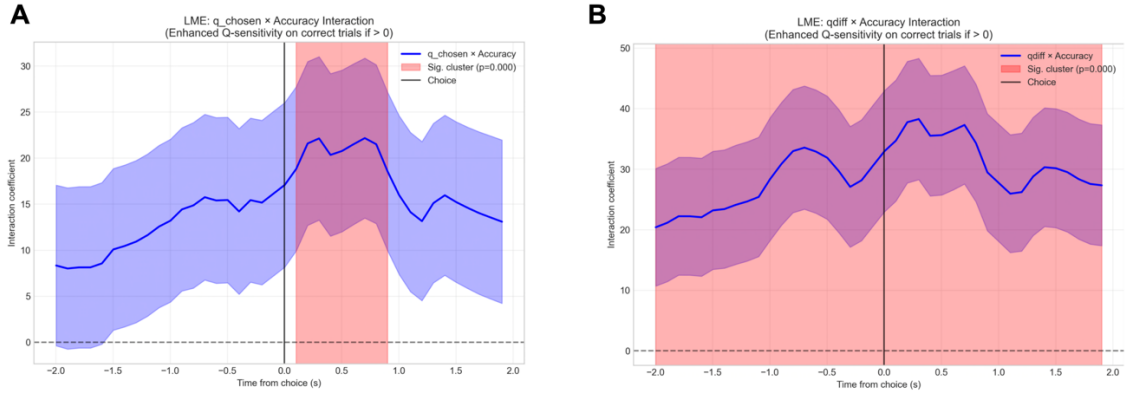


Figure 6. Enhanced Q-value sensitivity on accurate trials. A. Q_chosen × accuracy interaction showed enhanced sensitivity to chosen option value on correct trials during the post-choice period (0.1-0.9s, $p<0.001$). The positive interaction coefficient indicates stronger pupil-value coupling during outcome evaluation on accurate trials. **B.** Q_diff × accuracy interaction demonstrating enhanced sensitivity to Q_diff on correct trials across the entire window (-2.0-1.9s, $p<0.001$). Participants showing stronger pupil responses to value difference made more accurate choices, providing evidence for somatic marker function. Red shaded areas indicate significant interaction clusters after cluster-based permutation correction. Blue lines show interaction coefficients with 95% confidence intervals. The vertical line at 0s marks the choice onset.

Discussion

The present study provides compelling evidence for a refined understanding of the somatic marker hypothesis (SMH) in value-based decision-making under uncertainty. Using a novel combination of pupillometry, computational modelling, and probabilistic reinforcement learning, we demonstrated that physiological responses encode value differences during stimulus viewing and decision implementation through temporally distinct mechanisms, and that value-autonomic coupling during decision differs systematically between accurate and inaccurate trials.

Our findings reveal three key insights that collectively advance understanding of how bodily signals contribute to adaptive decision-making. First, stimulus-related pupil responses primarily encoded decision difficulty rather than the absolute reward magnitude of individual stimuli. The most robust finding was that pupil dilation correlated negatively with value differences between choice options (Q_diff) during stimulus evaluation, indicating that pupils dilate more when the value difference within the option pair is small. Second, choice-related pupil responses showed systematic correlations with learned value beliefs, with pupil responses tracking Q_chosen and Q_diff throughout the decision period. Notably, Q_diff showed a positive correlation during choice implementation, contrasting with the negative correlation during stimulus presentation. Third, and most critically for the SMH, pupil sensitivity to value information systematically differed between accurate and inaccurate trials. Trials resulting in correct choices demonstrated enhanced value-autonomic coupling spanning the whole decision interval, particularly for decision difficulty (Q_diff) measures.

These findings provide nuanced support for the SMH while refining our understanding of how somatic markers operate. The classic SMH proposed that anticipatory physiological responses signal the potential outcomes of different choices, biasing decisions toward advantageous options

even before conscious awareness (Bechara et al., 2005; Damasio, 1996). The enhanced Q-value sensitivity on accurate trials represents evidence for somatic marker function in our study, demonstrating that trials resulting in correct decisions are characterized by different patterns of value-autonomic coupling than those resulting in incorrect decisions. However, rather than providing binary signals about choice outcomes, our results suggest that somatic markers operate through differential sensitivity to value information. During stimulus evaluation, when the critical challenge is assessing relative values, larger pupil responses to small value differences (difficult choices) may mobilize cognitive resources for more careful evaluation. During choice implementation, when commitment is required, larger pupil responses to clear value differences may facilitate decisive action. This model reconciles seemingly contradictory findings in the literature and suggests that somatic markers are more sophisticated than previously theorized.

Our methodological approach addressed several limitations of previous SMH research. The use of pupillometry provided the temporal resolution necessary to isolate stimulus-specific responses, overcoming a major limitation of skin conductance measures. The sequential stimulus presentation design enabled within-trial comparison of physiological responses to competing options, while the computational modeling approach, which we adapted from Van Slooten et al., (2018), provided trial-by-trial estimates of value beliefs that captured individual learning trajectories rather than relying on objective reward probabilities, likely increasing our sensitivity to detect somatic marker effects. Our results particularly extend the work of Van Slooten et al., (2018), which demonstrated that pre-choice pupil dilation correlated with value beliefs and exploration-exploitation strategies, but did not directly test whether these responses differed systematically between correct and incorrect trials. Our interaction analyses provide the first direct

evidence that value-autonomic coupling differs between accurate and inaccurate decisions, supporting the functional significance of autonomic responses in value-based choice.

Several limitations warrant consideration. Our task design may not fully capture the complexity of real-world decision-making situations, and our sample of healthy young adults limits generalization to clinical populations or different age groups. A critical limitation concerns our failure to directly address the distinction between pre-aware and post-aware stages that is central to the original somatic marker hypothesis formulation. The classic IGT studies emphasized that somatic markers operate before conscious awareness of task contingencies, with anticipatory responses appearing even when participants could not explicitly articulate the reward structure. Our study did not systematically examine this awareness distinction, as we have not identified an optimal way to identify pre- versus post-aware stages. Additionally, we did not employ the deconvolution analysis approach used by Van Slooten et al., (2018), which was designed to disentangle overlapping physiological effects from temporally close experimental events. Our methodological choice may have conflated responses to different trial components, potentially obscuring specific mechanisms or introducing artifacts from overlapping physiological responses.

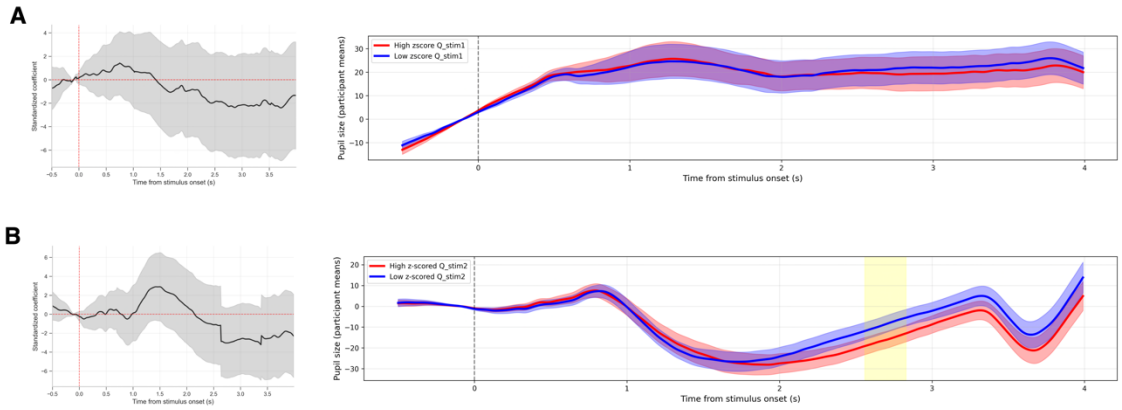
Alternative interpretations should be acknowledged. The enhanced value-autonomic coupling on accurate trials could reflect general factors like attention or motivation rather than specific somatic marker mechanisms. Participants who maintained better task engagement might show both stronger pupil-value correlations and better performance for reasons unrelated to somatic guidance. However, the specificity of our effects to value difference information and their presence before choice execution argues against purely attentional accounts. In addition, our focus on pupil dilation represents only one component of the autonomic nervous system, and it remains

unclear whether findings generalize to other autonomic measures such as skin conductance or heart rate variability that were central to the original somatic marker hypothesis studies.

Future research should pursue several directions. First, incorporating measures of conscious awareness would allow direct testing of whether pupil-value coupling emerges before explicit knowledge of task structure. Second, implementing deconvolution analyses to disentangle overlapping physiological responses would provide more precise characterization of the temporal dynamics of somatic marker function. Third, combining pupillometry with neuroimaging could identify neural substrates of value-autonomic integration, particularly in ventromedial prefrontal regions implicated in the original SMH. Lastly, extending this paradigm to clinical populations with known decision-making impairments would test the generalizability and potential diagnostic utility of our measures.

To conclude, our study provides evidence that pupil responses serve as reliable indicators of value-based decision-making processes, specifically encoding decision difficulty through temporally distinct mechanisms that differ between correct and incorrect trials. The differential coupling between physiological responses and value information on accurate versus inaccurate trials represents evidence that autonomic responses carry information about decision quality, supporting the fundamental premise that physiological signals are associated with adaptive behavior under uncertainty. However, our findings suggest that somatic markers operate through subtle modulation of value processing rather than providing clear-cut signals about choice outcomes. Our computational approach, adapted from Van Slooten et al., (2018), enabled us to examine these relationships using participants' value beliefs rather than objective probabilities, revealing the nuanced ways in which somatic signals contribute to decision-making.

Supporting information



S1 Fig. Stimulus-related pupil responses did not reflect value beliefs of individual stimulus (Q_stim). A. Left panel: linear mixed-effects results showed no significant correlation between Q_stim1 and stimulus1-related pupil responses. Right panel: within-subject median-split permutation showed no significant difference in stimulus-1-related pupil responses between high vs. low Q_stim1 groups. **B.** Left panel: linear mixed-effects results showed no significant correlation between Q_stim2 and stimulus2-related pupil responses. Right panel: within-subject median-split permutation presented a short significant cluster and, in general, showed no significant difference in stimulus-2-related pupil responses between high vs. low Q_stim2 groups.

References

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science (New York, N.Y.)*, *275*(5304), 1293–1295. https://doi.org/10.1126/science.275.5304.1293

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). The Iowa Gambling Task and the somatic marker hypothesis: Some questions and answers. *Trends in Cognitive Sciences*, *9*(4), 159–162. https://doi.org/10.1016/j.tics.2005.02.002

Becker, S., Bräscher, A.-K., Bannister, S., Bensafi, M., Calma-Birling, D., Chan, R. C. K., Eerola, T., Ellingsen, D.-M., Ferdenzi, C., Hanson, J. L., Joffily, M., Lidhar, N. K., Lowe, L. J., Martin, L. J., Musser, E. D., Noll-Hussong, M., Olino, T. M., Pintos Lobo, R., & Wang, Y. (2019). The role of hedonics in the Human Affectome. *Neuroscience and Biobehavioral Reviews*, *102*, 221–241. https://doi.org/10.1016/j.neubiorev.2019.05.003

Bierman, D., Cleeremans, A., van Ditzhuyzen, J., & Van Gaal, S. (2004). *The use of an implicit grammar task and eye measurements to study the somatic marker hypothesis*. Mimeo. Copy can be retrieved from: http://m0134. fmg. uva. nl/publications ….

Boersma, F., Wilton, K., Barham, R., & Muir, W. (1970). Effects of arithmetic problem difficulty on pupillary dilation in normals and educable retardates. *Journal of Experimental Child Psychology*, *9*(2), 142–155. https://doi.org/10.1016/0022-0965(70)90079-2

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, *76*, 1. https://doi.org/10.18637/jss.v076.i01

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology. General*, *143*(4), 1476–1488. https://doi.org/10.1037/a0035813

Chiu, Y.-C., Lin, C.-H., Huang, J.-T., Lin, S., Lee, P.-L., & Hsieh, J.-C. (2008). Immediate gain is long-term loss: Are there foresighted decision makers in the Iowa Gambling Task? *Behavioral and Brain Functions : BBF*, *4*, 13. https://doi.org/10.1186/1744-9081-4-13

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *351*(1346), 1413–1420. https://doi.org/10.1098/rstb.1996.0125

de Gee, J. W., Knapen, T., & Donner, T. H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(5), E618-625. https://doi.org/10.1073/pnas.1317557111

Guath, M., Willfors, C., Björlin Avdic, H., Nordgren, A., & Kleberg, J. L. (2023). Pupillary response in reward processing in adults with major depressive disorder in remission. *Journal of the International Neuropsychological Society: JINS*, *29*(3), 306–315. https://doi.org/10.1017/S1355617722000224

Guillaume, S., Jollant, F., Jaussent, I., Lawrence, N., Malafosse, A., & Courtet, P. (2009). Somatic markers and explicit knowledge are both involved in decision-making. *Neuropsychologia*, *47*(10), 2120–2124. https://doi.org/10.1016/j.neuropsychologia.2009.04.003

Hess, E. H., & Polt, J. M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science (New York, N.Y.)*, *143*(3611), 1190–1192. https://doi.org/10.1126/science.143.3611.1190

Kozunova, G. L., Sayfulina, K. E., Prokofyev, A. O., Medvedev, V. A., Rytikova, A. M., Stroganova, T. A., & Chernyshev, B. V. (2022). Pupil dilation and response slowing distinguish deliberate explorative choices in the probabilistic learning task. *Cognitive, Affective & Behavioral Neuroscience*, *22*(5), 1108–1129. https://doi.org/10.3758/s13415-022-00996-z

Lavín, C., San Martín, R., & Rosales Jubal, E. (2014). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience*, *7*, 218. https://doi.org/10.3389/fnbeh.2013.00218

Lempert, K. M., Chen, Y. L., & Fleming, S. M. (2015). Relating Pupil Dilation and Metacognitive Confidence during Auditory Decision-Making. *PLoS ONE*, *10*(5), e0126588. https://doi.org/10.1371/journal.pone.0126588

Lin, C.-H., Chiu, Y.-C., Lee, P.-L., & Hsieh, J.-C. (2007). Is deck B a disadvantageous deck in the Iowa Gambling Task? *Behavioral and Brain Functions*, *3*, 16. https://doi.org/10.1186/1744-9081-3-16

Maia, T. V., & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(45), 16075–16080. https://doi.org/10.1073/pnas.0406666101

Mckinnon, A. I., Gray, N. S., & Snowden, R. J. (2020). Enhanced emotional response to both negative and positive images in post-traumatic stress disorder: Evidence from

pupillometry. *Biological Psychology*, *154*, 107922. https://doi.org/10.1016/j.biopsycho.2020.107922

Simonovic, B., Stupple, E. J. N., Gale, M., & Sheffield, D. (2017, January 1). Pupil Dilation and Cognitive Reflection as Predictors of Performance on the Iowa Gambling Task. *CogSci 2017 - Proceedings of the 39th Annual Meeting of the Cognitive Science Society: Computational Foundations of Cognition*. https://repository.derby.ac.uk/item/937w9/pupil-dilation-and-cognitive-reflection-as-predictors-of-performance-on-the-iowa-gambling-task

Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., & Wagenmakers, E.-J. (2013). Performance of healthy participants on the Iowa Gambling Task. *Psychological Assessment*, *25*(1), 180–193. https://doi.org/10.1037/a0029929

Steinhauer, S. R., Siegle, G. J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology*, *52*(1), 77–86. https://doi.org/10.1016/j.ijpsycho.2003.12.005

Tomb, I., Hauser, M., Deldin, P., & Caramazza, A. (2002). Do somatic markers mediate decisions on the gambling task? *Nature Neuroscience*, *5*(11), 1103–1104; author reply 1104. https://doi.org/10.1038/nn1102-1103

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review*, *25*(6), 2005–2015. https://doi.org/10.3758/s13423-018-1432-y

Van Slooten, J. C., Jahfari, S., Knapen, T., & Theeuwes, J. (2018). How pupil responses track value-based decision-making during and after reinforcement learning. *PLoS Computational Biology*, *14*(11), e1006632. https://doi.org/10.1371/journal.pcbi.1006632

Wang, C.-A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., & Munoz, D. P. (2018). Arousal Effects on Pupil Size, Heart Rate, and Skin Conductance in an Emotional Face Task. *Frontiers in Neurology*, *9*. https://doi.org/10.3389/fneur.2018.01029

Xu, F., & Huang, L. (2020). Electrophysiological Measurement of Emotion and Somatic State Affecting Ambiguity Decision: Evidences From SCRs, ERPs, and HR. *Frontiers in Psychology*, *11*, 899. https://doi.org/10.3389/fpsyg.2020.00899