Associations with monetary values do not influence access to awareness for faces

Marcus Rothkirch [Corresp., 1] , **Maximilian Wieser** [1] , **Philipp Sterzer** [1, 2]

[1] Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

[2] Bernstein Center for Computational Neuroscience, Humboldt Universität Berlin, Berlin, Germany

Corresponding Author: Marcus Rothkirch
Email address: marcus.rothkirch@charite.de

Human faces can convey socially relevant information in various ways. Since the early detection of such information is crucial in social contexts, socially meaningful information might also have privileged access to awareness. This is indeed suggested by previous research using faces with emotional expressions. However, the social relevance of emotional faces is confounded with their physical stimulus characteristics. Here, we sought to overcome this problem by manipulating the relevance of face stimuli through classical conditioning: Participants had to learn the association between different face exemplars and high or low amounts of positive and negative monetary outcomes. Before and after the conditioning procedure, the time these faces needed to enter awareness was probed using continuous flash suppression, a variant of binocular rivalry. While participants successfully learned the association between the face stimuli and the respective monetary outcomes, faces with a high monetary value did not enter visual awareness faster than faces with a low monetary value after conditioning, neither for rewarding nor for aversive outcomes. Our results tentatively suggest that behaviorally relevant faces do not have privileged access to awareness when the assessment of the faces' relevance is dependent on the processing of face identity, as this requires complex stimulus processing that is likely limited at pre-conscious stages.

1 **Associations with monetary values do not influence**
2 **access to awareness for faces**
3

5 Marcus Rothkirch[1], Maximilian Wieser[1], Philipp Sterzer[1,2]
6
7 [1] Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, corporate
8 member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of
9 Health, Berlin, Germany
10 [2] Bernstein Center for Computational Neuroscience, Humboldt-Universität zu Berlin, Berlin,
11 Germany
12
13 Corresponding Author:
14 Marcus Rothkirch
15 Charitéplatz 1, Berlin, 10117, Germany
16 Email address: marcus.rothkirch@charite.de

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

Introduction

The ability to identify and to rapidly read information from human faces has a pivotal role in social contexts. Since the multitude of information conveyed by faces goes far beyond the image per se, different cognitive systems are involved in face processing. Faces are thus a popular tool to assess what types of information can be processed without the observer's awareness or have preferential access to awareness (Axelrod, Bar & Rees, 2015; Madipakkam & Rothkirch, 2019). A particular focus of previous research in this context was on the question whether the social meaning of faces is already processed at pre-conscious stages, thereby facilitating conscious awareness of faces that convey socially relevant information. Indeed, facial cues signaling threat (Yang, Zald & Blake, 2007; Yang & Yeh, 2018), trustworthiness (Stewart et al., 2012; Getov et al., 2015), or positive emotions (Stein & Sterzer, 2012) seem to accelerate the awareness of faces. However, emotional expressions and other social characteristics of faces, such as trustworthiness or dominance, are inextricably linked to physical stimulus properties, like contrast, luminance, or spatial frequencies. In fact, the prioritization of emotional faces can be largely explained by differences in such physical stimulus properties (Stein & Sterzer, 2012; Gray et al., 2013; Hedger et al., 2016; Stein et al., 2018). To be able to unequivocally attribute differences in the access to awareness of faces to their behavioral relevance, however, the influence of physical stimulus properties should be ruled out first (Moors et al., 2019).

An elegant way to circumvent the inherent confound between physical stimulus properties and higher-level relevance is to ascribe behavioral relevance to faces in a systematic and controlled manner. That way, the association between the physical characteristics and the relevance of stimuli can be balanced out across observers. Anderson et al. (2011) followed such an approach by pairing faces with positive, negative, or neutral gossip. They observed that faces previously paired with negative gossip dominated visual awareness during a following binocular rivalry task. In subsequent studies, in contrast, affective biographical information did not influence observers' awareness of faces, suggesting a rather limited impact of such information on visual awareness (Rabovsky, Stein & Abdel Rahman, 2016; Stein et al., 2017). One reason for these conflicting findings might be that the relevance of social information depends on each individual's evaluation of this information. Indeed, the time for complex stimuli to reach awareness can depend on the subjectively experienced value of the stimulus (Schmack et al., 2016) or certain personality traits of the observer (Madipakkam et al., 2019).

In the present study, we chose to systematically pair images of faces with high or low amounts of monetary reward and punishment. Manipulating the behavioral relevance of faces by means of monetary incentives has several advantages over using verbal descriptions. In comparison to the latter, monetary values are quantitative, which implies that different conditions can be clearly defined. In this regard, different conditions are set by different amounts of the same unit, whereas for biographical information descriptions of positive behaviors or traits, for example, are

78   compared to entirely different descriptions that are supposed to be identified as neutral. Thus,
79   monetary values allow for a systematic control of the behavioral relevance of faces and are
80   intersubjectively meaningful. The association of stimuli with monetary incentives by means of
81   classical conditioning can modulate the subjective value of the stimuli to a similar extent as to
82   primary reinforcers (Delgado, Labouliere & Phelps, 2006; Lehner et al., 2016). Moreover, such
83   associations have the potency to influence visual attention, such that stimuli associated with
84   higher monetary values capture and guide attention more strongly than stimuli associated with
85   lower values (Austin & Duka, 2010; Bucker & Theeuwes, 2017, 2018). It is reasonable to
86   assume that such learned associations can also affect how quickly stimuli enter awareness, since
87   stimuli previously paired with high monetary reward are more often consciously perceived
88   during rapid sequences of visual stimuli than stimuli paired with low reward (Leganes-
89   Fonteneau, Scott & Duka, 2018). In our study, participants had to learn the association between
90   face stimuli and monetary values. Before and after the learning phase, the same faces were
91   presented under breaking continuous flash suppression (bCFS) to measure the time they require
92   to get access to awareness. BCFS is a variant of binocular rivalry, in which two different stimuli
93   are presented to the two different eyes. Unlike binocular rivalry, however, for bCFS a dynamic
94   stimulus (i.e. the masking stimulus) is presented to one eye, while a static stimulus (i.e. the
95   target) is presented to the other eye. This results in the initial dominance of the masking stimulus,
96   allowing for a greater control of stimulus suppression in comparison to the rather stochastic
97   nature of perceptual states during binocular rivalry. We hypothesized that if learned values
98   facilitate faces' access to awareness, there should be a stronger decrease in response times after
99   the conditioning session for faces associated with a high monetary compared to a low monetary
100 value. The inclusion of monetary reward as well as punishment further enabled us to detect
101 valence-specific effects of learned values on visual awareness.
102
103

104 **Materials & Methods**
105

106 **Participants**
107 Twenty-four participants (13 females; age: 18 – 35 years, M = 24.54 ± 4.36 standard deviation
108 [SD]) took part in the experiment. Five participants were left-handed, all other participants were
109 right-handed. All participants had normal or corrected-to-normal vision and written informed
110 consent was obtained from each participant prior to their participation in the experiment. The
111 study was approved by the local ethics committee of the Charité – Universitätsmedizin Berlin
112 (EA1/301/13) and performed in accordance with the Declaration of Helsinki. The final sample
113 size was based on a sequential testing approach using Bayes Factors (Schönbrodt et al., 2017).
114 More specifically, we planned to calculate Bayes Factors (BF10) for our main analysis after
115 every new batch of 24 participants, since this sample size is needed to completely counterbalance
116 the association between stimulus exemplars and monetary values across participants. Our aim
117 was to continue data collection until these Bayes Factors would either exceed a threshold of 3,

118   which indicates evidence in favour of the alternative hypothesis, or fall below a threshold of 1/3,
119   indicating evidence in favour of the null hypothesis. Since this condition was already met after
120   the inclusion of 24 participants, we stopped data collection at this stage.
121
122   **Stimuli and Apparatus**
123   The face stimuli used in the study were grey-scale photographs of four different female faces
124   with a neutral expression, taken from the NimStim Set of Facial Expressions (Tottenham et al.,
125   2009; image IDs: 01, 07, 09, 17). All four images were similar in global contrast (root mean
126   square contrast between 0.16 and 0.20) and luminance (mean luminance between 27.49 cd/m²
127   and 31.35 cd/m²). All stimuli were presented on a uniformly grey background (30.28 cd/m²).
128   Participants viewed the screen through a mirror stereoscope providing separate visual input to
129   the two eyes. Each participant's head was stabilized by a chin rest at a viewing distance of 60
130   cm. All stimuli were presented using MATLAB (The MathWorks, Natick, MA, USA) and
131   Psychtoolbox-3 (http://psychtoolbox.org/) on a 19 inch CRT monitor (resolution: 1024 x 768 Px,
132   refresh rate: 60 Hz).
133
134   **Procedure**
135   The experiment consisted of three phases: 1) a pre-conditioning phase to measure baseline
136   response times, 2) a conditioning phase during which different faces were paired with monetary
137   outcomes, and 3) a post-conditioning phase that was identical to the pre-conditioning phase,
138   intended to assess the change in response times after the conditioning had taken place.
139   In the initial pre-conditioning phase (Figure 1A), different face exemplars were presented under
140   continuous flash suppression. At the beginning of each trial, a black rectangle (10° x 10°) and a
141   black fixation cross in its center (0.68° x 0.68°) were presented to each eye. The rectangle and
142   the cross were visible throughout the whole experimental phase. After a fixation duration of 2 s,
143   high-contrast dynamic mask stimuli consisting of circles and squares of various colors and sizes
144   were flashed to a randomly selected eye at a frequency of 10 Hz. Simultaneously, a face image
145   (3.75° x 3.75°) that was located within one of the four quadrants of the black rectangle was
146   presented to the other eye. The contrast of the face stimulus linearly increased from 0% to 100%
147   during the initial 2 s. After that, the face remained at full contrast until the end of the trial. A trial
148   ended when the participant gave a manual response or, if no response was made, after 15 s.
149   Participants' task was to indicate the location of the face, that is, the quadrant in which the face
150   appeared, as fast and as accurately as possible by pressing one of four designated keys on the
151   keyboard. This part of the experiment comprised 96 trials. The combination of the face
152   exemplar, the location of the face, and the eye to which the face was presented was
153   counterbalanced and randomized across trials.
154   In the second phase, participants had to learn the association between the face exemplars and
155   monetary outcomes by means of classical conditioning (Figure 1B). In each trial, one of the faces
156   that were already presented in the first part of the experiment was shown in the center of the
157   screen. Below the face, a positive or negative monetary value was displayed. The face and the

158   value were presented for 5 s. After the offset of the face and the associated monetary value, a
159   blank screen was presented for a randomized interval between 1 s and 1.6 s before the next trial
160   started. Each face was associated with one of four different monetary values: -2 €, -0.1 €, +0.1 €,
161   and +2 €. In 75% of the trials, the face was depicted with its associated monetary value. In the
162   remaining trials, an outcome of 0 € was presented along with the face. Such a probabilistic
163   reinforcement schedule was chosen to maintain participants' attention to the stimuli and the task,
164   as for classical conditioning with monetary outcomes attention is preferably directed towards
165   partially predictive stimuli in comparison to fully predictive stimuli (Austin & Duka, 2010).
166   Participants were instructed to passively view the stimuli and to memorize the association
167   between the faces and the monetary values as well as possible. They were further informed that
168   the monetary value depicted in a trial would be counted towards their overall payoff. Thus, for a
169   trial, in which a face was presented together with '+2 €', for instance, 2 € would be added to their
170   payoff, while for trials with '-2 €' 2 € would be subtracted from their payoff. Since unbeknown
171   to participants each monetary value was presented equally often, the outcome for this part of the
172   conditioning phase amounted to zero. Thus, participants' outcome was solely defined by their
173   accuracy in the query trials (see below). This phase of the experiment comprised four blocks,
174   during which each face was presented four times. The order of the face exemplars was
175   randomized and the association between the face exemplars and the monetary values was fully
176   counterbalanced across participants. The association was further kept constant across all four
177   blocks.
178   To assess whether participants were indeed able to learn these associations, query trials were
179   added at the end of each block. Each face was presented once during these query trials. In
180   contrast to the conditioning trials, a question mark and all four different monetary values were
181   displayed below the face in random order. Participants were required to select the value that was
182   associated with the respective face by pressing one of four designated buttons. The face and the
183   monetary values were presented until a response was made. Participants were informed that
184   monetary reimbursement for their participation in the experiment would depend on the accuracy
185   of their choices during these query trials. After the experiment, participants indeed received the
186   amount of money that they accumulated during the conditioning phase. More specifically, a
187   correct assignment of a monetary reward to the respective face would yield an addition of +2 €
188   or +0.1 €, respectively, to their payoff. A wrong response to faces associated with reward did not
189   result in a monetary gain. For faces associated with a monetary punishment, participants had to
190   assign the correct monetary value to avoid a monetary loss. This means that a correct assignment
191   of -2 € or -0.1 € to the respective face did not yield a monetary gain or loss. However, if a
192   participant assigned a wrong value to a punishment-related face, this resulted in a loss of -2 € or -
193   0.1 €, respectively. Thus, participants expected that their payoff was dependent on both, the
194   passively viewed pairings of monetary values and faces as well as their performance during the
195   query trials. Note that in case of negative values participants had to respond accurately to avoid
196   monetary losses during the query trials.

197   The post-conditioning phase, which followed directly after the conditioning phase, was fully
198   identical to the pre-conditioning phase. After the experiment, participants rated how much they
199   felt motivated by the different monetary values to memorize the faces on a visual analog scale
200   ranging from 0 to 5.
201
202   **Data Analysis**
203   Participants' learning performance during the conditioning phase was assessed on the basis of
204   their responses in the query trials. For each of the four blocks, each participant's accuracy in
205   assigning the monetary value to each face exemplar was computed. The chance level for each
206   block was .25. To evaluate participants' sensitivity for reward and punishment, we computed
207   each participant's response bias during the query trials. For reward-related faces the response
208   bias is computed on the basis of the z-transformed hits (H_R) and false alarm rates (FA_R) as
209   follows:

210    $$c_R = -\frac{z(H_R) + z(FA_R)}{2}$$

211   However, since participants have to assign a positive or negative value to each face, the response
212   biases for reward and punishment are not independent of each other. More specifically, an
213   increase in the false alarm rate for reward leads to a decrease in the hit rate for punishment (H_P)
214   by the same amount, such that: $z(FA_R) = -z(H_P)$. Thus, the formula above can be rewritten as:

215    $$c_R = -\frac{z(H_R) - z(H_P)}{2}$$

216   Since a bias towards reward (c_R) automatically implies a bias away from punishment (c_P) such
217   that $c_R = -c_P$, we only report the reward response bias. For c_R, a positive bias denotes a stronger
218   sensitivity for reward, while a negative bias denotes a stronger sensitivity for punishment.
219   For the bCFS phase before and after the conditioning phase, trials in which participants
220   responded incorrectly (percentage of trials: M = 3.19%, SD = 2.75%) or failed to give a response
221   until the end of a trial (M = 0.61%, SD = 1.26%) were discarded from further analysis.
222   Furthermore, response times below 200 ms were considered anticipatory responses and were also
223   discarded from further analysis (percentage of trials: M = 0.61%, SD = 1.26%). To compare the
224   response times between the different conditions, we followed two approaches. For the first
225   approach, we computed the median response time for each participant and condition before and
226   after conditioning. We used the median instead of the mean to account for the skewness of
227   response time distributions, which is line with other bCFS studies (e.g. Gayet et al., 2016; Gayet
228   & Stein, 2017; Han, Blake & Alais, 2018). We then subtracted the median response times of the
229   pre-conditioning phase from the median response times of the post-conditioning phase, which
230   resulted in a measure for the change of response times for each condition and participant.
231   Finally, we performed two paired t-tests to compare the mean change in response times between
232   high and low reward as well as between high and low punishment. The alpha level of these two
233   t-tests was adjusted to .025 to account for multiple comparisons. For the second approach, we
234   analyzed the response time distributions in more depth by computing hierarchical shift functions

235 (Rousselet & Wilcox, 2019), which can be more sensitive to response time differences,
236 especially when they are restricted to early or late responses. To this end, we computed the
237 deciles of the response time distribution for each condition and participant before and after
238 conditioning. In a next step, we subtracted the deciles of the pre-conditioning phase from the
239 deciles of the post-conditioning phase. The resulting values thus indicate the change in response
240 times for each segment of the whole distribution, where lower deciles reflect faster responses and
241 higher deciles slower responses. For each decile, we then performed a paired t-test to compare
242 the changes in response times of the high reward to the low reward condition and of the high
243 punishment to the low punishment condition. As this amounts to 18 different t-tests in total, we
244 adjusted the alpha level of the t-tests to .003. Thus, we adjusted the significance level for the
245 number of tests across reward and punishment, because our aim was to control the maximum
246 experiment-wise error rate (Bender & Lange, 2001). Our reasoning for applying such a rigorous
247 adjustment of the alpha level was that the underlying hypothesis of an influence of learned
248 values on bCFS response times would be supported by a difference in response times for either
249 reward or punishment.
250 To probe potential time-dependent effects of learned values, we additionally split the post-
251 conditioning phase into two halves. For statistical inference, we performed repeated-measures
252 ANOVAs with the factors time (first half vs. second half of the post-conditioning phase) and
253 value (high value vs. low value). Separate ANOVAs were performed for reward-related and
254 punishment-related stimuli. The focus of this analysis was on the interaction between time and
255 value, since a statistically significant interaction would signify an influence of the monetary
256 value that was dependent on time. Note that this analysis was performed on the median response
257 time differences between the pre- and post-conditioning phase.
258 Finally, we performed two generalized linear mixed-effects models on the single trial data, one
259 for reward-related and one for punishment-related stimuli. This analysis was intended to take
260 into account the variability of the effect of interest in dependence on the repetition of the stimuli
261 and the association between the different face exemplars and the monetary values. We defined a
262 Gamma distribution as the distribution of our response variable, since it provides a plausible
263 approximation to the processes reflected in response times (Lo & Andrews, 2015). We included
264 the interaction of face value, bCFS phase, face exemplar, and trial number as well as all other
265 interactions and main effects of these variables as fixed effects in our models, for which we used
266 effects coding. For the random effects term, we defined the theoretically maximal model by
267 including subject as a random factor together with the by-subject random intercept and the by-
268 subject random slopes of all fixed factors (Barr et al., 2013). Both models thus had the following
269 structure:
270 $RT \sim value * phase * exemplar * trial + (1 + value * phase * exemplar * trial | subject)$
271 The estimate of each fixed-effects predictor was then tested against zero by means of individual
272 t-tests.
273

274   In addition to frequentist inference statistics, we computed Bayes Factors (BF10) for each t-test,
275   using a default Cauchy prior of 0.707. For ANOVAs we computed BF10 directly from the F-
276   values of the ANOVA statistics (Faulkenberry, 2018). In line with previous suggestions (Wetzels
277   & Wagenmakers, 2012), we interpret BF10 > 3 as evidence for the alternative hypothesis and
278   BF10 < 1/3 as evidence for the null hypothesis. We also report BF01 to clarify which analyses
279   yielded evidence for the null hypothesis. In this case, BF01 > 3 indicates evidence for the null
280   hypothesis.

281

282

283   **Results**

284

285   **Classical conditioning**
286   At the end of every block in the conditioning phase, participants had to indicate the associated
287   monetary value for each presented face. Figure 2A depicts the accuracy of these responses for
288   the four different blocks. Since participants had to make four choices in each block, the chance
289   level corresponds to an accuracy of .25 for each block. Binomial tests indicated that the response
290   accuracies across all four blocks exceeded chance level for all participants (p ≤ .002, BF10 ≥
291   10.10). Figure 2B shows the response accuracies across all four blocks separately for each value
292   condition. As for each value condition the majority of participants did not commit any mistake
293   during the query trials, the median accuracy for each condition amounted to 1. Thus, participants
294   quickly and successfully learned the associations between the face exemplars and the monetary
295   outcomes for each of the different monetary values. After the experiment, participants rated their
296   motivation for each monetary value on a visual analog scale ranging from 0 to 5. In comparison
297   to a low reward of 0.1 € (M = 1.51 ± 0.28 standard error of the mean [SEM]), participants felt
298   substantially more motivated by the high reward of 2 € (M = 2.98 ± 0.36 SEM; paired t-test:
299   t(23) = 4.99, p < .001, BF10 = 521.53). The difference in motivation between a low punishment
300   of 0.1 € (M = 1.46 ± 0.30 SEM) and a high punishment of 2 € (M = 2.08 ± 0.37 SEM) was less
301   pronounced compared to reward (t(23) = 2.42, p = .024, BF10 = 2.35), but still statistically
302   significant at a corrected threshold of α = .025. The majority of participants (n = 18) did not
303   exhibit a response bias to either reward or punishment (i.e. c_R = 0). One participant showed a
304   response bias towards monetary punishment (c_R = -0.18) and five participants showed a response
305   bias towards monetary reward (c_R = 0.24 – 0.97).

306

307   **Response times during breaking-CFS**
308   Figure 3A shows the change in response times from the pre-conditioning to the post-conditioning
309   phase for the two different reward conditions. For the high reward condition, response times
310   decreased by 501.01 ms (± 102.27 ms SEM), on average, while for the low reward condition we
311   observed an average decrease in response times of 412.74 ms (± 84.25 ms SEM). The difference
312   between the two conditions was not statistically significant (t(23) = 0.80, p = .43) and the Bayes
313   factor indicated the absence of an effect (BF10 = 0.29, BF01 = 3.48).

314   The average change in response times for the high and low punishment condition are depicted in
315   Figure 3B. Numerically, there was a stronger decrease in response times for the low punishment
316   condition (M = -477.85 ± 97.54 ms SEM) in comparison to the high punishment condition (M =
317   -401.97 ± 82.05 ms SEM). However, the difference between the two conditions was again not
318   statistically significant and indicative of an absence of an effect (t(23) = 0.65, p = .52, BF10 =
319   0.26, BF01 = 3.84).
320   To rule out baseline differences between conditions that might have masked potential effects of
321   learned values in the post-conditioning phase, we additionally computed response times during
322   the pre-conditioning phase only. Response times for faces that were later paired with a high
323   reward (M = 2275.90 ± 464.57 ms SEM) did not significantly differ from response times for
324   faces that were later paired with a low reward (M = 2254.57 ± 460.21 ms SEM; t(23) = 0.17, p =
325   .87, BF10 = 0.22, BF01 = 4.60). Similarly, the time to respond to faces later associated with high
326   punishment (M = 2224.52 ± 454.08 ms SEM) was not significantly different from faces later
327   associated with low punishment (M = 2315.77 ± 472.71 ms SEM; t(23) = 0.85, p = .40, BF10 =
328   0.30, BF01 = 3.36).
329
330   Differences between high and low values, however, might be limited to a specific range of the
331   whole response time distributions, that is, the learned values could specifically affect fast or slow
332   responses, which would not necessarily be captured by measures of central tendencies. We
333   therefore explored the response time distributions in more depth by computing the change
334   between the pre- and the post-conditioning phase for each decile of the whole response time
335   distribution for each condition. For reward-related stimuli (Figure 4A) as well as well as for
336   punishment-related stimuli (Figure 4B), the decrease in response times was more pronounced for
337   slower responses. However, there was no significant difference between high and low values for
338   any of the deciles (p ≥ .12 and BF10 ≤ 0.65 for reward, p ≥ .12 and BF10 ≤ 1.66 for punishment).
339   Of note, the differences between high and low reward and punishment, respectively, were close
340   to an uncorrected significance level of .05 for the last deciles. The response time decreases in this
341   decile, however, were numerically stronger for low values compared to high values and as such
342   contrary to our a priori hypothesis.
343
344   Due to extinction, potential influences of the monetary associations on response times could have
345   quickly decayed during the post-conditioning phase. Thus, averaging across all responses in the
346   post-conditioning phase might have masked the effects of learned values. To identify potential
347   time-dependent effects, we split the post-conditioning phase into two halves. We performed two
348   repeated-measures ANOVAs with the factors time (first half vs. second half of the post-
349   conditioning phase) and value (low value vs. high value). If the influence of learned values was
350   indeed dependent on time such that it quickly faded after the conditioning block, this should be
351   reflected in the interaction between time and value. However, neither for faces previously
352   associated with reward (F(1,23) = 0.60, p = .44, BF10 = 0.28, BF01 = 3.59) nor for faces

353   associated with punishment ($F(1,23) = 0.76$, $p = .39$, $BF10 = 0.30$, $BF01 = 3.31$) we observed an
354   interaction between time and value.

355   In line with our main analysis, the parameter estimates for the interaction of the face value and
356   the bCFS phase in the generalized linear mixed-effects models were not statistically significant
357   from zero, neither for reward ($\beta = -8.49 \times 10^{-6}$, $t(2194) = -1.87 \times 10^{-5}$, $p > .99$) nor for punishment
358   ($\beta = -4.26 \times 10^{-5}$, $t(2193) = -8.62 \times 10^{-5}$, $p > .99$). There were also no other statistically significant
359   interactions with trial number or face exemplar (all $p > .99$).

360

361   *Exploratory analyses*
362   While our a priori hypotheses focused on the comparison of the change in response times for
363   high vs. low monetary values, we performed two additional exploratory analyses to examine
364   whether the learned values might have influenced the response times during the bCFS phase in
365   other ways.
366   For the first analysis, we collapsed our data across high and low values and compared response
367   time changes for all reward-related faces to all punishment-related faces. For reward-related
368   faces, response times decreased by 428.13 ms (± 87.39 SEM), while for punishment-related
369   faces response times decreased by 443.98 ms (± 90.63 SEM), on average. The difference
370   between reward and punishment was not statistically significant ($t(23) = 0.19$, $p = .85$, $BF10 =$
371   0.22, $BF01 = 4.59$). Even when only faces associated with a high value (i.e. high reward vs. high
372   punishment) were considered, the difference between reward and punishment did not reach
373   statistical significance ($t(23) = 1.15$, $p = .26$, $BF10 = 0.39$, $BF01 = 2.58$).
374   In a second exploratory analysis we examined whether the influence of learned values on
375   response times might have been dependent on the learning performance during the conditioning
376   phase. To this end, we computed each participant's accuracy in the query trials across all four
377   conditioning blocks. We then correlated these accuracies against the response time differences
378   (i.e. post- vs. pre-conditioning) during the bCFS blocks, separately for reward- and punishment-
379   related faces. The bCFS response time differences were computed by subtracting the change in
380   response times for the low value condition from the change in response times for the high value
381   condition as follows: ($[RT_{post} - RT_{pre}]_{high\ reward} - [RT_{post} - RT_{pre}]_{low\ reward}$) and ($[RT_{post} - RT_{pre}]_{high}$
382    $_{punishment} - [RT_{post} - RT_{pre}]_{low\ punishment}$). As such, negative values indicate a change in response
383   times in line with our a priori hypotheses, that is, a stronger decrease in response times for the
384   high compared to the low value condition. Since only few participants exhibited low accuracies
385   during the query trials, the distribution of participants' accuracies was heavily skewed. To
386   account for this, we computed Spearman's ρ. Neither for reward ($\rho = .30$, $p = .16$, $BF10 = 0.43$,
387    $BF01 = 2.35$) nor for punishment ($\rho = .13$, $p = .56$, $BF10 = 0.19$, $BF01 = 5.39$), accuracies were
388   related to response time differences between conditions. Note that if higher accuracies in the
389   query trials were related to a greater influence of high monetary values on response times, this
390   would be indicated by a negative correlation. Due to the high accuracy during the query trials in
391   the majority of participants, this analysis should be interpreted with caution.

392

393

Discussion

395

396 We studied whether learned values, established by means of classical conditioning with
397 monetary reward and punishment, influence access to awareness for faces. While participants
398 successfully learned the association between the different face exemplars and the monetary
399 values, the learned association did not have an influence on their response times. Response times
400 generally decreased from the pre- to the post-conditioning phase. However, this decrease was
401 equally strong for high compared to low reward and for high compared to low punishment. A
402 more in-depth exploration of the response time distributions did not reveal an advantage for faces
403 paired with a higher compared to a lower monetary value either.

404

405 Faces can express and convey their relevance in various ways, for instance through their
406 emotional expression or particular facial features such as eye gaze. In the current study, we
407 paired faces with monetary incentives to render them behaviorally relevant. This way, we
408 intended to circumvent a confound between physical stimulus characteristics and higher-level
409 social relevance, which is especially prevalent in the investigation of the awareness of emotional
410 faces (Hedger et al., 2016). While it has previously been suggested that learned associations
411 between affective information and faces affect the faces' potency to dominate awareness
412 (Anderson et al., 2011), subsequent studies did not observe a privileged access to awareness for
413 faces paired with affective information (Rabovsky, Stein & Abdel Rahman, 2016; Stein et al.,
414 2017). Our findings are in line with the latter studies, while at the same time complementing
415 these previous results by showing that not only biographical information but also monetary
416 incentives fail to facilitate awareness of faces. The approach of associating face stimuli with
417 affective information, however, differs in two aspects from the investigation of the influence of
418 emotional expressions. First, emotional expressions are inherent to a face, while the association
419 with affective information has to be learned. Secondly, the identification of emotional
420 expressions often only requires the processing of certain facial features. A rapid detection of
421 fearful faces, for instance, is likely due to the greater exposure to the iris and the sclera in fearful
422 faces (Whalen et al., 2004). Effects of learned associations, in contrast, likely requires the
423 identification of the faces' whole identity. Since conditioned responses to fear-conditioned faces
424 transfer to novel images of the same face identity (Rehbein et al., 2018), learned associations are
425 indeed, at least partly, related to the face's identity. Thus, the influence of learned affective
426 information, as in our case by means of monetary values, is dependent on a more complex
427 analysis of the stimulus at pre-conscious stages. As the processing of face identity is rather
428 limited under visual masking (Moradi, Koch & Shimojo, 2005; Amihai, Deouell & Bentin,
429 2011), the scope of pre-conscious face processing might not be sufficient to boost faces into
430 awareness that have been coupled with positive or negative outcomes. Similar limitations likely
431 apply to other types of complex stimuli when different exemplars of the same category are paired

432   with different outcomes. Though to what extent pre-conscious processing is limited for different
433   stimulus categories has to be further investigated in future studies.
434
435   Participants were clearly able to learn the association between the face exemplars and the
436   respective monetary values. The absence of an influence of the associated values on the access of
437   the faces to awareness can thus not be attributed to a failure or difficulties of participants to learn
438   these associations. There is ample evidence for sustained neural (Rothkirch et al., 2012) and
439   behavioral effects (Raymond & O'Brien, 2009; Rutherford, O'Brien & Raymond, 2010;
440   Rothkirch et al., 2013) of previously learned associations between faces and monetary values,
441   indicating that pairing face stimuli with monetary outcomes has the potency to render face
442   stimuli behaviorally relevant for an extended period of time. It must be noted, however, that such
443   associations have been mostly induced by means of instrumental conditioning so far. While
444   classical conditioning with monetary incentives can generally bring about similar effects
445   compared to instrumental conditioning (Delgado, Labouliere & Phelps, 2006; Bucker &
446   Theeuwes, 2016, 2017), the specific combination of face stimuli and monetary outcomes in the
447   context of classical conditioning has only rarely been studied so far. However, Trilla Gros,
448   Panasiti & Chakrabarti (2015) found that EEG responses to faces are influenced by their
449   previous associations with monetary reward and punishment established through classical
450   conditioning. While this shows that classical conditioning with monetary values has the potency
451   to alter neural signals in response to face stimuli, it leaves open the question how such altered
452   neural signals translate into behavioral effects. It has further been demonstrated previously that
453   simple visual stimuli, like gratings, can gain faster access to awareness by means of classical fear
454   conditioning (Gayet et al., 2016), showing that a conditioning approach can, in principle, confer
455   behavioral relevance to visual stimuli such that they access awareness more rapidly. There is the
456   residual possibility that the association between the face stimuli and the monetary values in our
457   study did not effectively change the affective content of the faces, especially because associating
458   faces with monetary incentives by means of classical conditioning might be less effective
459   compared to instrumental conditioning. For fear-conditioning, the effectiveness of the
460   conditioning procedure is usually assessed on the basis of physiological measures, like skin
461   conductance responses. For appetitive conditioning and conditioning with monetary outcomes, in
462   contrast, such a standard physiological measure has not yet been established. Since pupil size
463   promises to be a fruitful measure of the effectiveness of appetitive conditioning (Pietrock et al.,
464   2019), it could be assessed in future studies focusing on the influence of learned stimulus values
465   on the access to awareness. The availability of such a measure would also allow to relate the
466   strength of conditioning in each individual to the effect of learned values on visual awareness
467   (Madipakkam et al., 2016; Vieira et al., 2017). Alternatively or complementary to physiological
468   responses, the effectiveness of the conditioning procedure could further be evaluated on the basis
469   of a control task with clearly visible stimuli. Such a control task, though, would either require a
470   separate sample of participants or would have to be included after the conditioning phase in
471   addition to the bCFS task, in which case the extinction of the conditioned values plays an

472   important role. Physiological measures, in contrast, could be more easily integrated in the
473   existing experimental design.
474   A further relevant aspect are potential asymmetries in participants' sensitivity to reward or
475   punishment. We assessed participants' sensitivity in our study on the basis of their responses
476   during the query trials in the conditioning phase. Overall, we observed a slightly greater
477   sensitivity towards reward compared to punishment, suggesting that for some participants
478   seeking rewards was more relevant during the conditioning phase than avoiding punishment.
479   However, the overwhelming majority of participants did not exhibit a bias in any direction and
480   almost all participants were able to correctly assign the monetary values to the different face
481   exemplars until the end of the conditioning phase. Notably, in case of negative values
482   participants had to respond accurately to avoid monetary losses during the query trials. It is
483   possible that this task has rendered the loss-related faces similar to the reward-related faces in
484   terms of their motivational content. It has indeed been argued previously that the avoidance of
485   aversive outcomes can be rewarding (Kim, Shimojo & O'Doherty, 2006), which would imply
486   that the loss-related faces in our study might have been perceived as motivationally positive.
487   This assumption is in conflict with several other studies, however, indicating that separate neural
488   structures underlie reward and avoidance learning (Yacubian et al., 2006; Palminteri et al., 2012;
489   Rothkirch et al., 2017), which suggests that the avoidance of punishment is qualitatively distinct
490   from receiving a reward.
491   For the assessment of possible asymmetries in the sensitivity between reward and punishment or
492   motivational similarities between seeking rewards and avoiding losses, it is important to take into
493   account that we compared response times in the bCFS phase between high and low values
494   separately for reward and punishment. Thus, any influences that might have affected the relation
495   between reward- and punishment-related faces could not have biased our results.
496
497   It is conceivable that an influence of learned values on response times might have been
498   dampened either due to habituation because of the repeated exposure to each face exemplar or
499   due to extinction during the post-conditioning phase. Furthermore, it has previously been
500   reported that responses to fear-conditioned faces rapidly decrease when these faces are visually
501   masked (Raio et al., 2012). Thus, effects of learned values can be obscured when responses are
502   aggregated across the post-conditioning phase. However, an analysis that distinguished between
503   the first and second half of the post-conditioning phase did not provide any indication of such
504   time-dependent effects in our study. Moreover, while Raio et al. (2012) conducted the
505   conditioning procedure with faces that were suppressed from awareness, which likely established
506   only unstable associations between the conditioned and unconditioned stimuli, the conditioning
507   procedure in our study was performed with fully visible face stimuli. Still, each face exemplar
508   was repeated 64 times across all phases in our study. Indeed, in comparison to previous studies
509   (Anderson et al., 2011; Rabovsky, Stein & Abdel Rahman, 2016; Stein et al., 2017) our
510   conditioning phase comprised more repetitions of each face exemplar and a lower number of
511   different face exemplars. It has to be noted, however, that this ensured that participants

512   successfully learned the associations between faces and values. Rabovsky, Stein & Abdel
513   Rahman (2016), in contrast, report that in their study only 36% of newly learned associations
514   could be explicitly recalled after the learning phase, on average. In the other two studies
515   (Anderson et al., 2011; Stein et al., 2017), the learning procedure was repeated until participants
516   reached a criterion of at least 60% correct responses. Consequently, participants differed in the
517   frequency with which they were exposed to the face stimuli. Furthermore, they underwent the
518   post-conditioning task even though some of them might not have learned the correct associations
519   for a substantial amount of faces until the end of the learning phase. This suggests that fewer
520   repetitions during the learning phase likely come at the expense of a poorer learning
521   performance, which could have also contributed to the discrepant findings in previous studies.
522   According to Vansteenwegen et al. (2006), the affective content that faces gain through classical
523   conditioning is further largely resistant to extinction, at least at the behavioral level. In their first
524   experiment, response times in an affective priming task were still influenced by the learned
525   values after extinction, even though the experiment comprised 60 repetitions of each of the two
526   different conditioned face stimuli in total, which is comparable to the 64 repetitions of each face
527   exemplar in our study.
528
529   While we used faces with a neutral expression in our study, a potential approach to further
530   strengthen the association between the faces and the monetary values is to use faces with an
531   emotional expression. As suggested by the 'preparedness hypothesis', faces with different
532   emotional expressions might be differentially prepared to become associated with different
533   outcomes (Dimberg & Öhman, 1996). In this context, pairing aversive outcomes with angry
534   faces, for instance, might be more effective than pairing them with neutral faces. The specific
535   interactions between different emotional expressions and monetary outcomes have not been
536   systematically studied yet, however, and such an approach my come at the expense of potential
537   ceiling effects (Lonsdorf et al., 2017). Finally, while we have used monetary outcomes to render
538   face stimuli behaviorally relevant, the use of other reinforcers, like liquid rewards in water-
539   deprived participants or bursts of white noise, are conceivable alternatives. We chose monetary
540   outcomes as they are easy to administer and can be equally employed as rewarding and aversive
541   stimuli. Furthermore, the processing of primary and secondary reinforcers, including monetary
542   values, shows large overlaps in the human brain (Izuma, Saito & Sadato, 2008; Delgado, Jou &
543   Phelps, 2011; Sescousse et al., 2013), which suggests that monetary values can evoke similar
544   positive or negative experiences in comparison to other types of reinforcers.
545
546
547   **Conclusions**
548
549   To conclude, we did not observe a privileged access to awareness for faces that were associated
550   with positive or negative monetary outcomes, although participants quickly learned these
551   associations. This tentatively suggests that learned values that are tied to a face's identity have

552 only limited influence on the face's access to awareness, as such an influence possibly exceeds
553 the scope of pre-conscious processing.
554
555

556 **References**

557

558 Amihai I, Deouell L, Bentin S. 2011. Conscious awareness is necessary for processing race and
559   gender information from faces. *Consciousness and Cognition* 20:269–279. DOI:
560   10.1016/j.concog.2010.08.004.
561 Anderson E, Siegel EH, Bliss-Moreau E, Barrett LF. 2011. The visual impact of gossip. *Science*
562   *(New York, N.Y.)* 332:1446–1448. DOI: 10.1126/science.1201574.
563 Austin AJ, Duka T. 2010. Mechanisms of attention for appetitive and aversive outcomes in
564   Pavlovian conditioning. *Behavioural Brain Research* 213:19–26. DOI:
565   10.1016/j.bbr.2010.04.019.
566 Axelrod V, Bar M, Rees G. 2015. Exploring the unconscious using faces. *Trends in Cognitive*
567   *Sciences* 19:35–45. DOI: 10.1016/j.tics.2014.11.003.
568 Barr DJ, Levy R, Scheepers C, Tily HJ. 2013. Random effects structure for confirmatory
569   hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68:255–278.
570   DOI: 10.1016/j.jml.2012.11.001.
571 Bender R, Lange S. 2001. Adjusting for multiple testing--when and how? *Journal of Clinical*
572   *Epidemiology* 54:343–349. DOI: 10.1016/s0895-4356(00)00314-0.
573 Bucker B, Theeuwes J. 2016. Appetitive and aversive outcome associations modulate
574   exogenous cueing. *Attention, Perception & Psychophysics* 78:2253–2265. DOI:
575   10.3758/s13414-016-1107-6.
576 Bucker B, Theeuwes J. 2017. Pavlovian reward learning underlies value driven attentional
577   capture. *Attention, Perception & Psychophysics* 79:415–428. DOI: 10.3758/s13414-016-
578   1241-1.
579 Bucker B, Theeuwes J. 2018. Stimulus-driven and goal-driven effects on Pavlovian associative
580   reward learning. *Visual Cognition* 26:131–148. DOI: 10.1080/13506285.2017.1399948.
581 Delgado MR, Jou RL, Phelps EA. 2011. Neural systems underlying aversive conditioning in
582   humans with primary and secondary reinforcers. *Frontiers in Neuroscience* 5:71. DOI:
583   10.3389/fnins.2011.00071.
584 Delgado MR, Labouliere CD, Phelps EA. 2006. Fear of losing money? Aversive conditioning
585   with secondary reinforcers. *Social Cognitive and Affective Neuroscience* 1:250–259.
586   DOI: 10.1093/scan/nsl025.
587 Dimberg U, Öhman A. 1996. Behold the wrath: Psychophysiological responses to facial stimuli.
588   *Motivation and Emotion* 20:149–182. DOI: 10.1007/BF02253869.
589 Gayet S, Paffen CLE, Belopolsky AV, Theeuwes J, Van der Stigchel S. 2016. Visual input
590   signaling threat gains preferential access to awareness in a breaking continuous flash
591   suppression paradigm. *Cognition* 149:77–83. DOI: 10.1016/j.cognition.2016.01.009.
592 Gayet S, Stein T. 2017. Between-Subject Variability in the Breaking Continuous Flash
593   Suppression Paradigm: Potential Causes, Consequences, and Solutions. *Frontiers in*
594   *Psychology* 8:437. DOI: 10.3389/fpsyg.2017.00437.

595 Getov S, Kanai R, Bahrami B, Rees G. 2015. Human brain structure predicts individual
596   differences in preconscious evaluation of facial dominance and trustworthiness. *Social*
597   *Cognitive and Affective Neuroscience* 10:690–699. DOI: 10.1093/scan/nsu103.
598 Gray KLH, Adams WJ, Hedger N, Newton KE, Garner M. 2013. Faces and awareness: low-
599   level, not emotional factors determine perceptual dominance. *Emotion (Washington,*
600   *D.C.)* 13:537–544. DOI: 10.1037/a0031403.
601 Han S, Blake R, Alais D. 2018. Slow and steady, not fast and furious: Slow temporal modulation
602   strengthens continuous flash suppression. *Consciousness and Cognition* 58:10–19. DOI:
603   10.1016/j.concog.2017.12.007.
604 Hedger N, Gray KLH, Garner M, Adams WJ. 2016. Are visual threats prioritized without
605   awareness? A critical review and meta-analysis involving 3 behavioral paradigms and
606   2696 observers. *Psychological Bulletin* 142:934–968. DOI: 10.1037/bul0000054.
607 Izuma K, Saito DN, Sadato N. 2008. Processing of social and monetary rewards in the human
608   striatum. *Neuron* 58:284–294. DOI: 10.1016/j.neuron.2008.03.020.
609 Kim H, Shimojo S, O'Doherty JP. 2006. Is Avoiding an Aversive Outcome Rewarding? Neural
610   Substrates of Avoidance Learning in the Human Brain. *PLOS Biology* 4:e233. DOI:
611   10.1371/journal.pbio.0040233.
612 Leganes-Fonteneau M, Scott R, Duka T. 2018. Attentional responses to stimuli associated with
613   a reward can occur in the absence of knowledge of their predictive values. *Behavioural*
614   *Brain Research* 341:26–36. DOI: 10.1016/j.bbr.2017.12.015.
615 Lehner R, Balsters JH, Herger A, Hare TA, Wenderoth N. 2016. Monetary, Food, and Social
616   Rewards Induce Similar Pavlovian-to-Instrumental Transfer Effects. *Frontiers in*
617   *Behavioral Neuroscience* 10:247. DOI: 10.3389/fnbeh.2016.00247.
618 Lo S, Andrews S. 2015. To transform or not to transform: using generalized linear mixed models
619   to analyse reaction time data. *Frontiers in Psychology* 6. DOI:
620   10.3389/fpsyg.2015.01171.
621 Lonsdorf TB, Menz MM, Andreatta M, Fullana MA, Golkar A, Haaker J, Heitland I, Hermann A,
622   Kuhn M, Kruse O, Meir Drexler S, Meulders A, Nees F, Pittig A, Richter J, Römer S,
623   Shiban Y, Schmitz A, Straube B, Vervliet B, Wendt J, Baas JMP, Merz CJ. 2017. Don't
624   fear "fear conditioning": Methodological considerations for the design and analysis of
625   studies on human fear acquisition, extinction, and return of fear. *Neuroscience and*
626   *Biobehavioral Reviews* 77:247–285. DOI: 10.1016/j.neubiorev.2017.02.026.
627 Madipakkam AR, Rothkirch M. 2019. The Unconscious Processing of Social Information. In:
628   *Transitions Between Consciousness and Unconsciousness*. DOI:
629   10.4324/9780429469688-4.
630 Madipakkam AR, Rothkirch M, Dziobek I, Sterzer P. 2019. Access to awareness of direct gaze
631   is related to autistic traits. *Psychological Medicine* 49:980–986. DOI:
632   10.1017/S0033291718001630.
633 Madipakkam AR, Rothkirch M, Wilbertz G, Sterzer P. 2016. Probing the influence of
634   unconscious fear-conditioned visual stimuli on eye movements. *Consciousness and*
635   *Cognition* 46:60–70. DOI: 10.1016/j.concog.2016.09.016.
636 Moors P, Gayet S, Hedger N, Stein T, Sterzer P, van Ee R, Wagemans J, Hesselmann G. 2019.
637   Three Criteria for Evaluating High-Level Processing in Continuous Flash Suppression.
638   *Trends in Cognitive Sciences* 23:267–269. DOI: 10.1016/j.tics.2019.01.008.

639   Moradi F, Koch C, Shimojo S. 2005. Face adaptation depends on seeing the face. *Neuron*
640   45:169–175. DOI: 10.1016/j.neuron.2004.12.018.
641   Palminteri S, Justo D, Jauffret C, Pavlicek B, Dauta A, Delmaire C, Czernecki V, Karachi C,
642   Capelle L, Durr A, Pessiglione M. 2012. Critical Roles for Anterior Insula and Dorsal
643   Striatum in Punishment-Based Avoidance Learning. *Neuron* 76:998–1009. DOI:
644   10.1016/j.neuron.2012.10.017.
645   Pietrock C, Ebrahimi C, Katthagen TM, Koch SP, Heinz A, Rothkirch M, Schlagenhauf F. 2019.
646   Pupil dilation as an implicit measure of appetitive Pavlovian learning. *Psychophysiology*
647   56:e13463. DOI: 10.1111/psyp.13463.
648   Rabovsky M, Stein T, Abdel Rahman R. 2016. Access to Awareness for Faces during
649   Continuous Flash Suppression Is Not Modulated by Affective Knowledge. *PloS One*
650   11:e0150931. DOI: 10.1371/journal.pone.0150931.
651   Raio CM, Carmel D, Carrasco M, Phelps EA. 2012. Nonconscious fear is quickly acquired but
652   swiftly forgotten. *Current biology: CB* 22:R477-479. DOI: 10.1016/j.cub.2012.04.023.
653   Raymond JE, O'Brien JL. 2009. Selective visual attention and motivation: the consequences of
654   value learning in an attentional blink task. *Psychological Science* 20:981–988. DOI:
655   10.1111/j.1467-9280.2009.02391.x.
656   Rehbein MA, Pastor MC, Moltó J, Poy R, López-Penadés R, Junghöfer M. 2018. Identity and
657   expression processing during classical conditioning with faces. *Psychophysiology*
658   55:e13203. DOI: 10.1111/psyp.13203.
659   Rothkirch M, Ostendorf F, Sax A-L, Sterzer P. 2013. The influence of motivational salience on
660   saccade latencies. *Experimental Brain Research* 224:35–47. DOI: 10.1007/s00221-012-
661   3284-4.
662   Rothkirch M, Schmack K, Schlagenhauf F, Sterzer P. 2012. Implicit motivational value and
663   salience are processed in distinct areas of orbitofrontal cortex. *NeuroImage* 62:1717–
664   1725. DOI: 10.1016/j.neuroimage.2012.06.016.
665   Rothkirch M, Tonn J, Köhler S, Sterzer P. 2017. Neural mechanisms of reinforcement learning
666   in unmedicated patients with major depressive disorder. *Brain: A Journal of Neurology*
667   140:1147–1157. DOI: 10.1093/brain/awx025.
668   Rousselet GA, Wilcox RR. 2019. *Reaction times and other skewed distributions: problems with*
669   *the mean and the median*. PsyArXiv. DOI: 10.31234/osf.io/3y54r.
670   Rutherford HJV, O'Brien JL, Raymond JE. 2010. Value associations of irrelevant stimuli modify
671   rapid visual orienting. *Psychonomic Bulletin & Review* 17:536–542. DOI:
672   10.3758/PBR.17.4.536.
673   Schmack K, Burk J, Haynes J-D, Sterzer P. 2016. Predicting Subjective Affective Salience from
674   Cortical Responses to Invisible Object Stimuli. *Cerebral Cortex* 26:3453–3460. DOI:
675   10.1093/cercor/bhv174.
676   Schönbrodt FD, Wagenmakers E-J, Zehetleitner M, Perugini M. 2017. Sequential hypothesis
677   testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*
678   22:322–339. DOI: 10.1037/met0000061.
679   Sescousse G, Caldú X, Segura B, Dreher J-C. 2013. Processing of primary and secondary
680   rewards: a quantitative meta-analysis and review of human functional neuroimaging
681   studies. *Neuroscience and Biobehavioral Reviews* 37:681–696. DOI:
682   10.1016/j.neubiorev.2013.02.002.

683   Stein T, Awad D, Gayet S, Peelen MV. 2018. Unconscious processing of facial dominance: The
684   role of low-level factors in access to awareness. *Journal of Experimental Psychology.*
685   *General* 147:e1–e13. DOI: 10.1037/xge0000521.
686   Stein T, Grubb C, Bertrand M, Suh SM, Verosky SC. 2017. No impact of affective person
687   knowledge on visual awareness: Evidence from binocular rivalry and continuous flash
688   suppression. *Emotion (Washington, D.C.)* 17:1199–1207. DOI: 10.1037/emo0000305.
689   Stein T, Sterzer P. 2012. Not just another face in the crowd: detecting emotional schematic
690   faces during continuous flash suppression. *Emotion (Washington, D.C.)* 12:988–996.
691   DOI: 10.1037/a0026944.
692   Stewart LH, Ajina S, Getov S, Bahrami B, Todorov A, Rees G. 2012. Unconscious evaluation of
693   faces on social dimensions. *Journal of Experimental Psychology. General* 141:715–727.
694   DOI: 10.1037/a0027950.
695   Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Marcus DJ, Westerlund A,
696   Casey BJ, Nelson C. 2009. The NimStim set of facial expressions: judgments from
697   untrained research participants. *Psychiatry Research* 168:242–249. DOI:
698   10.1016/j.psychres.2008.05.006.
699   Trilla Gros I, Panasiti MS, Chakrabarti B. 2015. The plasticity of the mirror system: how reward
700   learning modulates cortical motor simulation of others. *Neuropsychologia* 70:255–262.
701   DOI: 10.1016/j.neuropsychologia.2015.02.033.
702   Vansteenwegen D, Francken G, Vervliet B, De Clercq A, Eelen P. 2006. Resistance to
703   extinction in evaluative conditioning. *Journal of Experimental Psychology. Animal*
704   *Behavior Processes* 32:71–79. DOI: 10.1037/0097-7403.32.1.71.
705   Vieira JB, Wen S, Oliver LD, Mitchell DGV. 2017. Enhanced conscious processing and
706   blindsight-like detection of fear-conditioned stimuli under continuous flash suppression.
707   *Experimental Brain Research* 235:3333–3344. DOI: 10.1007/s00221-017-5064-7.
708   Whalen PJ, Kagan J, Cook RG, Davis FC, Kim H, Polis S, McLaren DG, Somerville LH, McLean
709   AA, Maxwell JS, Johnstone T. 2004. Human amygdala responsivity to masked fearful
710   eye whites. *Science (New York, N.Y.)* 306:2061. DOI: 10.1126/science.1103617.
711   Yacubian J, Gläscher J, Schroeder K, Sommer T, Braus DF, Büchel C. 2006. Dissociable
712   Systems for Gain- and Loss-Related Value Predictions and Errors of Prediction in the
713   Human Brain. *Journal of Neuroscience* 26:9530–9537. DOI: 10.1523/JNEUROSCI.2915-
714   06.2006.
715   Yang Y-H, Yeh S-L. 2018. Can emotional content be extracted under interocular suppression?
716   *PloS One* 13:e0206799. DOI: 10.1371/journal.pone.0206799.
717   Yang E, Zald DH, Blake R. 2007. Fearful expressions gain preferential access to awareness
718   during continuous flash suppression. *Emotion (Washington, D.C.)* 7:882–886. DOI:
719   10.1037/1528-3542.7.4.882.
720

Figure 1

Schematic depiction of the experimental procedure.

(A) In the pre- and post-conditioning phase, high-contrast dynamic mask stimuli were presented to one eye at a frequency of 10 Hz. A face stimulus was simultaneously presented to the other eye. The contrast of the face stimulus linearly increased during the initial 2 s and remained at full contrast until the end of the trial. Participants' task was to indicate the location of the face. A trial ended either after a manual response or at the latest after 15 s. **(B)** In the conditioning phase, different face stimuli were presented together with their associated monetary outcome (upper part in panel B). Participants' task was to passively view and memorize these associations. This phase comprised four blocks. At the end of each block, query trials where performed (lower part in panel B), where each face stimulus was presented with four different response options. Here, participants had to select the value that was previously associated with the face. Due to copyright reasons, the original face images used in the experiment are not shown in the figure but replaced by a schematic drawing.
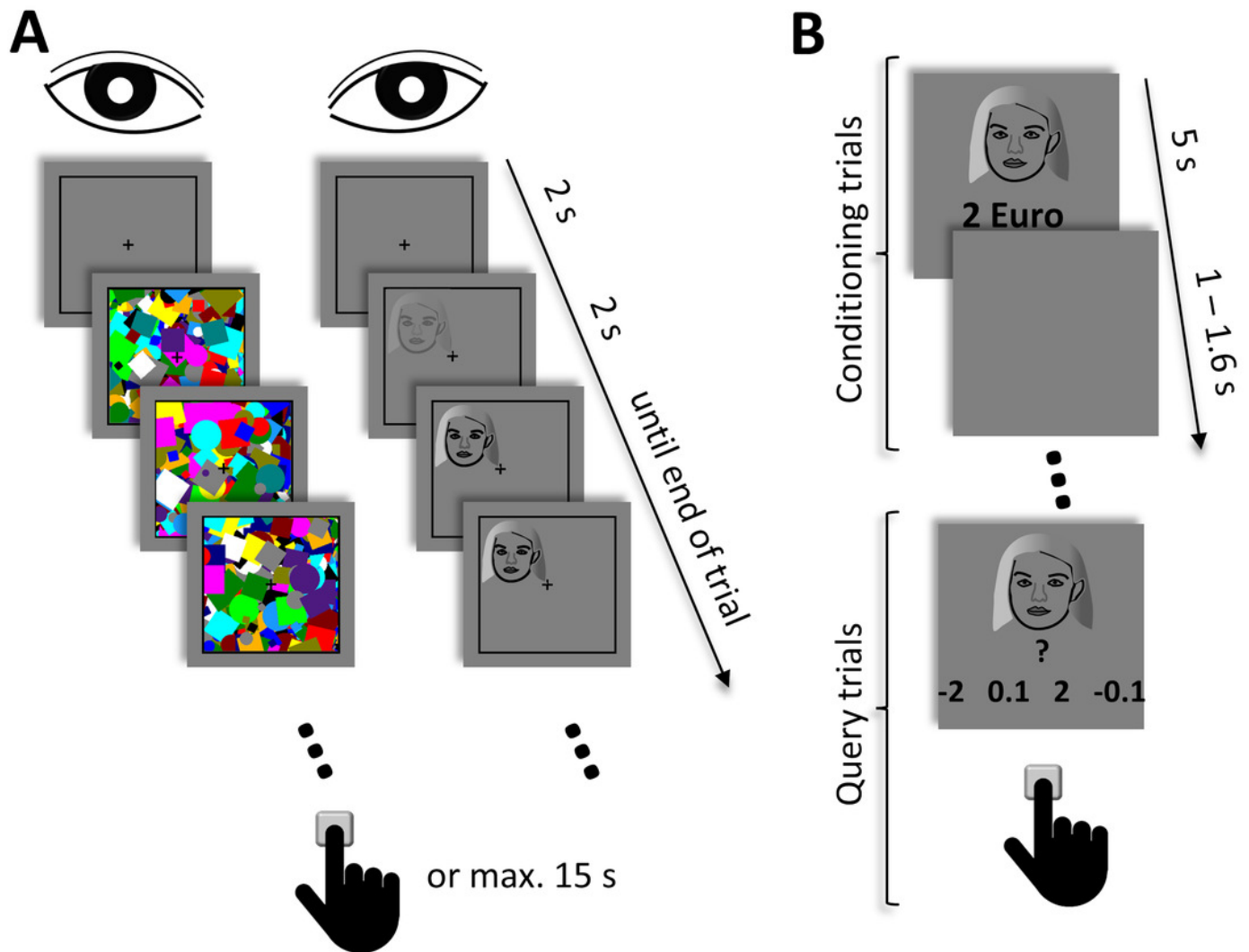
Figure 2

Choice accuracy in the query trials of the conditioning phase.

(A) Accuracy for each block of the conditioning phase. **(B)** Accuracy for each value condition.
Each dot represents one participant.
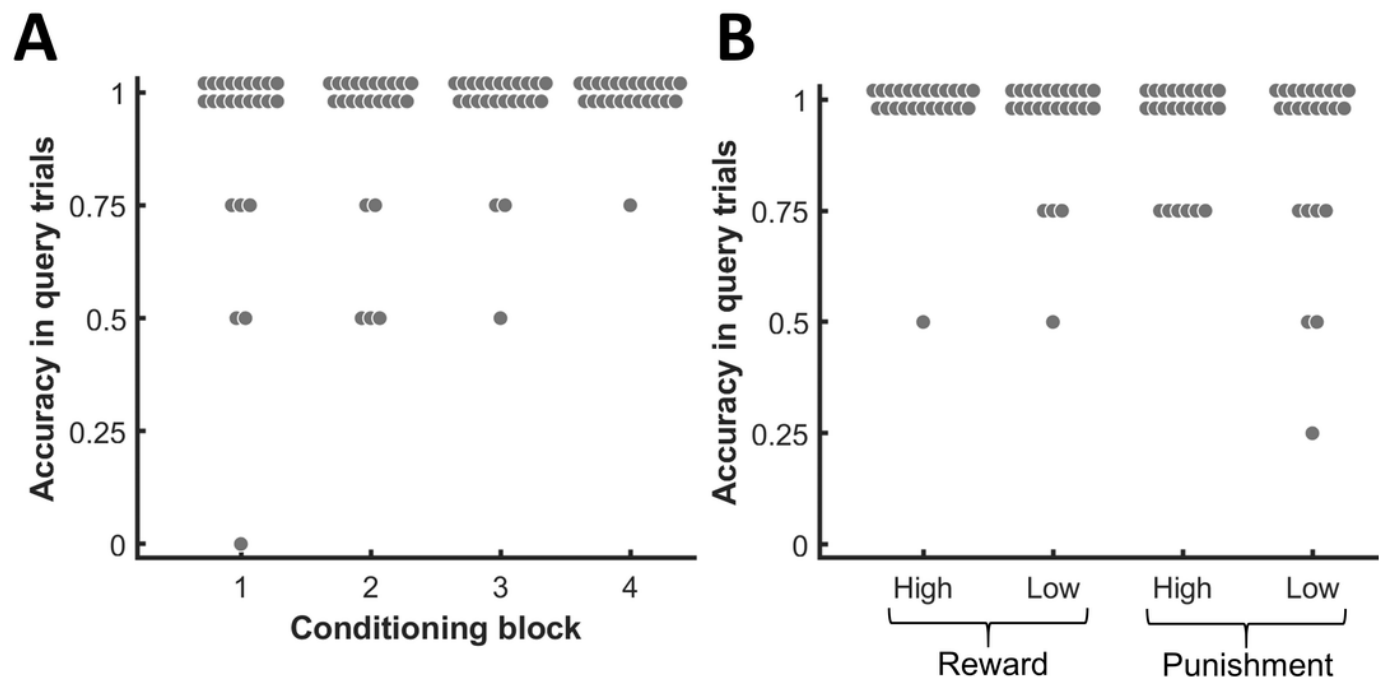
Figure 3

Change in response times (RT) for (A) the reward-related and (B) the punishment-related stimuli.

Negative values indicate an RT decrease in the post-conditioning phase compared to the pre-conditioning phase. The position of each data point is defined by the RT change for the high value condition on the x-axis and the RT change for the low value condition on the y-axis. Participants who showed a stronger RT reduction for the high in comparison to the low value condition (i.e. RT differences consistent with our a priori hypotheses) are located in the grey-shaded area. The magenta-coloured dashed line indicates the mean RT change for the high value condition across participants. The turquoise dashed line indicates the mean RT change for the low value condition across participants.
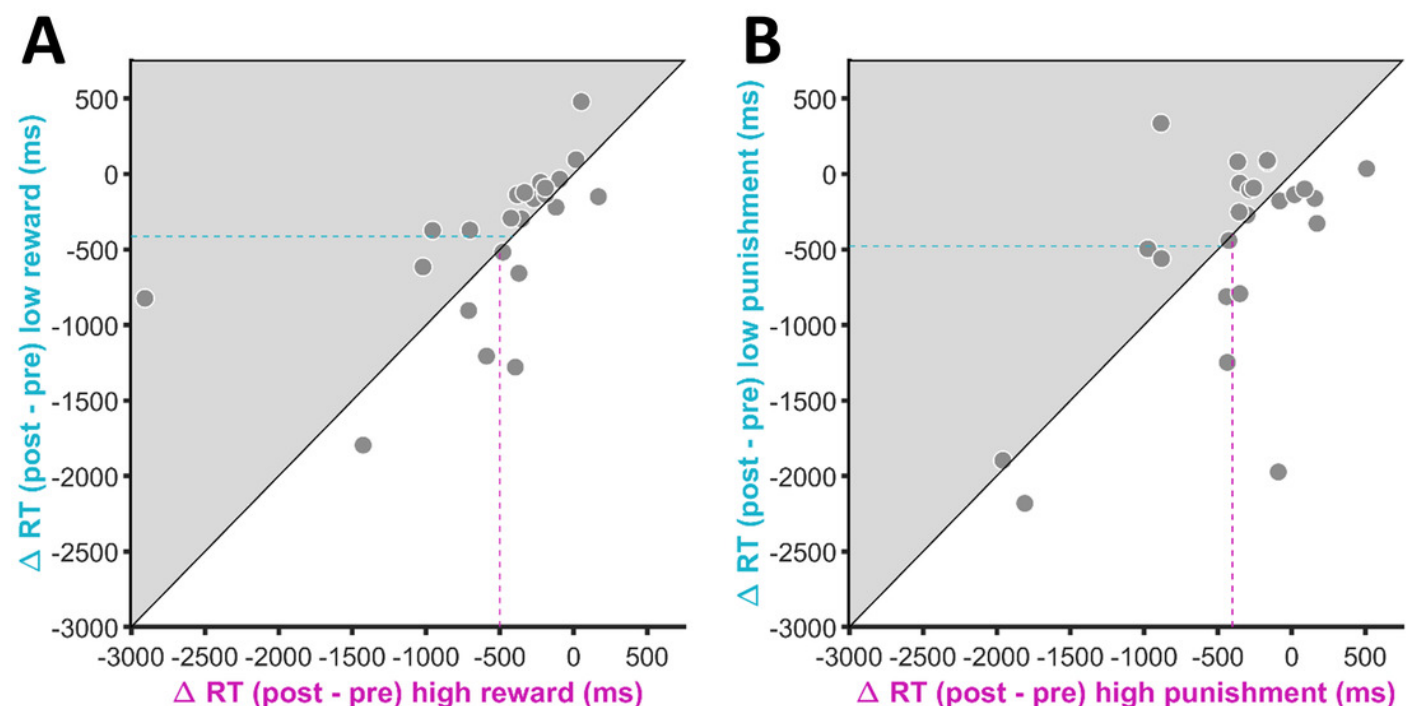
Figure 4

Change in response times for each decile of the whole response distributions for the (A) reward and (B) punishment conditions.

The numbers at the top of each graph indicate the p-values and the Bayes Factors for the comparison between the two conditions for each respective decile. The error bars display standard errors of the mean.