Cogn Comput (2017) 9:457–467
DOI 10.1007/s12559-017-9456-6

CrossMark

Neuronal Network and Awareness Measures of Post-Decision Wagering Behavior in Detecting Masked Emotional Faces

Remigiusz Szczepanowski [1] · **Michał Wierzchoń** [2] · **Marcin Szulżycki** [1]

Received: 16 March 2016 / Accepted: 21 February 2017 / Published online: 7 March 2017

Abstract Awareness can be measured by investigating the patterns of associations between discrimination performance (first-order decisions) and confidence judgments (knowledge). In a typical post-decision wagering (PDW) task, participants judge their performance by wagering on each decision made in a detection task. If participants are aware, they wager advantageously by betting high whenever decisions are correct and low for incorrect decisions. Thus, PDW—like other awareness measures with confidence ratings—quantifies if the knowledge upon which they make their decisions is conscious. The present study proposes a new method of assessing the association between advantageous wagering and awareness in the PDW task with a combination of log-linear (LLM) modeling and neural network simulation to reveal the computational patterns that establish this association. We applied the post-decision wagering measure to a backward masking experiment in which participants made first-order decisions about whether or not a masked emotional face was present, and then used imaginary or real monetary stakes to judge the correctness of their initial decisions. The LLM analysis was then used to examine whether advantageous wagering was aware by testing a hypothesis of partial associations between metacognitive judgments and accuracy of first-order decisions. The LLM outcomes were submitted into a feed-forward neural network. The network served as a general approximator that was trained to learn relationships between input wagers and the output of the corresponding log-linear function. The simulation resulted in a simple network architecture that successfully accounted for wagering behavior. This was a feed-forward network unit consisting of one hidden neuron layer with four inputs and one output. In addition, the study indicated no effect of the monetary incentive cues on wagering strategies, although we observed that only low-wager input weights of the neural network considerably contributed to advantageous wagering.

✉ Remigiusz Szczepanowski
rszczepanowski@swps.edu.pl

[1] Wroclaw Faculty of Psychology, SWPS University of Social Sciences and Humanities in Wroclaw, Wroclaw, Poland

[2] Consciousness Lab, Institute of Psychology, Jagiellonian University, Krakow, Poland

Keywords Awareness · Metacognition · Connectionist model · Post-decision wagering · Log-linear analysis

Introduction

There are several experimental methods for investigating conscious and unconscious processing of visual stimuli (see e.g., [9, 25, 39]). For instance, given the visibility threshold, the salience of the stimulus or its presentation times manipulated under a backward masking task enables dissociations between conscious and unconscious perception [14, 29]. Nevertheless, behavioral experiments in which a participant systematically uses metacognitive judgments (e.g., confidence ratings) to assess the accuracy of perceptual discriminations often encounter difficulty in estimating the subjective threshold of awareness (see e.g., [33]). For example, a post-decision wagering paradigm (see below) that rests on the assumption that advantageous wagering indicates awareness may be subjected to a variety of confounding factors such as priming effects, motivation, and loss aversion. Therefore, it is important for consciousness research to develop a measurement method that is sensitive enough to reveal the subjective threshold. Here, we propose a novel method based on log-linear modeling (LLM) that determines metacognitive awareness in terms of the accuracy of first-order discriminations. We also examine the computational grounds underlying metacognitive awareness using a neural network that captures the associations between wagers and the accuracy of first-order discriminations.

The aforementioned post-decision wagering paradigm has been extensively used to estimate metacognitive awareness in perception [26, 28, 33]. Under a typical post-decision wagering task, participants are asked to discriminate a stimulus (e.g., male vs. female face, fearful or neutral facial expression) and then are required to express their confidence in the discrimination decisions using monetary stakes. Hence, participants aim to maximize possible profits so that they use advantageous wagers when they are aware of the accuracy of first-order discriminations [13, 26]. It is often argued that post-decision wagering does not require introspection ability; therefore, it seems to be more natural and intuitive for participants than other subjective measures of awareness (see [26]). Despite the advantages of PDW, this method has been also criticized in the literature. For instance, participants use high bets when they are fully aware of the stimuli, but it is less evident what levels of awareness are associated with lower wagers [33]. This is mainly due to specific response strategies such as loss aversion (see [31]) that can affect results collected with the PDW scale. For this reason, some researchers argue that post-decision wagering may be less sensitive to low-level awareness [11, 33, 37]. In fact, the loss aversion strategy results in low wager ratings even when participants are aware of the stimuli (but not confident in the decision), and thus, the lowest ratings do not reflect unconscious processing (guessing criterion [10]). Moreover, the awareness ratings collected with post-decision wagering are also modified by the monetary values of the available stakes [13].

It is important to note that post-decision wagering may be also influenced by task-induced motivation, which, in turn, can affect either first-order discriminations (see [35]) or metacognitive awareness (see [38]). In fact, several studies show that task-induced motivation may be further manipulated by the usage of real or artificial stakes while wagering (see e.g., [2, 11, 26]). Some researchers argue that wagering with imaginary incentives is as effective as betting with real money (see [26]). In contrast, Dienes and Seth [11] suggest that wagering with real money could induce higher motivation in task performance. These researchers also argue that PDW may be a less sensitive measure of awareness during anticipation of loss of money. Taking these effects together, this clearly shows that post-decision wagering may be subjected to several confounding factors that need to be accounted for when estimating the subjective threshold of metacognitive awareness.

Another important issue is how to simulate advantageous wagering and awareness. Accessing conscious evaluation requires the presence of first-order discriminations; nevertheless, it strictly depends on knowledge that corresponds to such first-order information [4]. To address the problem of metacognitive knowledge formation, wagering behavior can be simulated with a neural network. In fact, neural networks have been proven to be effective in simulating first-order discriminations such as facial perception (see e.g., [8, 17, 22, 30]). Similarly, a simulation study with a neural network by Cleeremans et al. [6] demonstrated that the connectionist approach could be useful in simulating metacognitive awareness in post-decision wagering tasks. These researchers found that a hierarchical neural network was able to reproduce behavioral results and could acquire knowledge about its own behavior (see [6, 24, 36]). Importantly, the neural network simulation by Cleeremans et al. [6] was not focused on response strategies applied to wagering and therefore could not provide clear assessments of the individual threshold of awareness. This was mainly due to the fact that the neural network assumed the behavior of an ideal participant who showed a correlation between awareness and advantageous wagering strategy, but, in fact, was not influenced by additional confounds, for example loss aversion.

It is important to note that Szczepanowski [33] has shown that assessment of confounding factors on wagering performance should be focused on responses associated with the lowest wagers. In particular, LLM modeling demonstrated that a typical examination of advantageous wagering fails to measure the zero-accuracy criterion [33]. In particular, LLM analysis of partial associations between accuracy and wagering indicated that some people exhibit conscious knowledge even though advantageous wagering was absent. Therefore, to assess whether a participant is conscious, researchers should examine advantageous wagering by studying partial associations among proportions of low wagers (for more details, see the “Method” section). In fact, given that conscious evaluation of performance is independent from expression of behavior, a partial association counterpart of advantageous wagering can be further simulated with a neural network, (see [4, 6]). Note that during the process of learning, the neural network predicts the next element of a sequence in the task to be trained, and in this fashion, it acquires some form of knowledge [24] which is then stored and encoded within its architecture and input weights. Thus, from a computational perspective, the neuronal network in some sense establishes preconditions for knowledge that enables classifications between aware vs. unaware states while wagering on the accuracy of the first-order discrimination [24].

To sum up, in our study, we behaviorally tested participants with a post-wagering task at two target durations (i.e., 17- and 33-ms targets) and assumed that these conditions allowed us to compare aware and unaware wagering behavior. To manipulate the sensory threshold, we employed a backward masking task with emotional faces. This follows a vast body of empirical reports [27, 34] demonstrating that this technique is efficient for establishing subliminal perception [18, 23], especially for fearful stimuli presented at near-threshold for durations less than 25 ms (see [34]). Thus, in our study, the initial first-order decisions were made upon fearful facial expressions, and then imaginary or real monetary stakes were used to reveal metacognitive awareness of correct or incorrect first-order discriminations. To make sure that the measures of awareness we provide are accurate, we estimated the subjective threshold with a metacognitive sensitivity measure based on a "type 2" signal detection theory that also is thought to be independent of response biases [21]. Next, we simulated behavior with a neural network that presumably encoded different input weights associated with aware and unaware wagering. We hypothesized that a simulated wagering strategy should reflect behavior that is more effective when the network wagered for the aware condition (33-ms targets) as opposed to the unaware condition (17 ms), for which higher wagers are more likely to follow incorrect first-order classification.

Method

Experiment

In order to distinguish post-decision wagering judgment strategies in aware and unaware conditions, we applied a subliminal representation of visual stimuli by engaging subjects in a backward masking task involving briefly presented, facial expression targets that were subsequently impaired by a mask presentation. In this fashion, we could measure their sensory consciousness once a threshold for the facial expression was established below which participants reported perceiving the prime [5]. Subsequently, participants were required to wager a small amount of money to reveal conscious knowledge about their perception. We hypothesized that under the backward masking condition a longer presentation of the emotional face should result in a more effective wagering strategy as indexed by monetary prize winnings, as opposed to the unaware condition under which subliminal presentation should be reflected by less effective first-order discrimination that also leads the neural network to less effective classifications.

Participants

Twenty-six students (19 females) of the Wroclaw Faculty of Psychology, SWPS School of Social Sciences and Humanities in Wroclaw participated in this study. Their age ranged from 19 to 37, with an average of 23. The research was approved by the Ethics Committee of SWPS University of Social Sciences and Humanities. All subjects had normal or corrected-to-normal vision. The population sample was randomly assigned to both groups in terms of incentive treatment (real vs. imaginary monetary incentives), i.e., half the participants wagered with real money, while the second group wagered with imaginary money. Due to failures to perform the wagering task, datasets from two participants were removed from the sample. Instructions for the "real money" group led subjects to believe that winnings were dependent on task performance; however, all participants were given maximum of 50 PLN at the end of the experiment.

Stimuli and Procedure

Participants were seated in a darkened room and their heads were positioned on a chinrest to fix their position and reduce head movement. The stimuli were displayed in the centre of an Iiyama MA203DT Vision Master Pro 513 monitor at a screen refresh rate of 120 Hz, driven by an ATI Radeon HD 4800 Series graphics card.

In each trial, participants saw a green fixation cross (300 ms), then a blank screen (50 ms) and a target face, depicting either fearful or neutral facial expressions (17- or 33-ms face target) (see Fig. 1). The facial target was followed by a neutral face, which was a mask. The total duration of the target-mask pair was established at 133 ms, resulting in 100- or 117-ms durations of the mask. After presentation of the masked target, participants had to discriminate emotion, indicating whether the target face was fearful or not, and then were asked to rate their decision awareness by selecting one of two coins (1 or 2 PLN) on the visual display as monetary wagers; the decision periods were limited to 2 and 2.5 s, respectively. Participants started with no winnings and subsequently won the wagered amount of money; however, when they gambled wrongly, they lost that amount. No feedback on actual winnings was provided throughout the procedure. Both responses were recorded using a numerical response pad. After a 500-ms inter-trial interval, a new trial started. The total duration of each trial was 5 s. Trials were presented in a random order, with half the trials containing fearful-neutral target-mask pairs and the other half containing neutral-neutral target-mask pairs. Forty fearful and 40 neutral faces were selected as targets and an additional 80 neutral face stimuli were used as masks. We used images from three sets of faces: the Paul Ekman set [12], the set described by Öhman and colleagues (KDEF, [19]), and a third set validated by Alumit Ishai at NIMH (Bethesda,

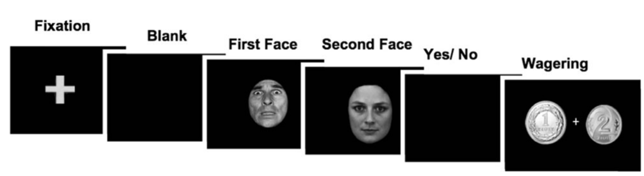

Fig. 1 Experimental procedure of post-decision wagering with masked faces. At the beginning of each trial, a green fixation cross was presented at the center of the computer screen for 300 ms. A blank screen followed for 50 ms. The target face depicting either a fearful or a neutral expression with equal probability was displayed at the central position for 17- or 33-ms and was immediately masked by a neutral face. The total duration of the target-mask pair was fixed at 133 ms, resulting in 100- or 117-ms exposures to the mask. After presentation of the target-mask pair, participants had 2 s to determine whether the target face was fearful or neutral, and were given an additional 2.5 s to express confidence with low and high wagers (1 PLN or 2 PLN)

USA; [16]). Each stimulus subtended 4° × 5° of the visual angle and was presented in the center of the computer screen at a viewing distance of 50 cm. Participants performed the masking task in two experimental conditions (17- and 33-ms targets), divided into eight blocks (40 trials each). The order of blocks was randomized across the experiment.

Awareness Measure Using Log-Linear Modeling

It has been argued that post-decision wagering demonstrates that perception can occur without conscious awareness when wagering is independent of task performance (for the correctness of the first-order discriminations, see [26]). Szczepanowski [33] indicated that the LLM method is useful for examining whether participants are aware or unaware of having information about the correctness of their decisions. In particular, the log-linear approach accepts that a null hypothesis needs only partial association between low wagers to be tested. Thus, to examine the hypothesis of no partial association between awareness and advantageous wagering, a special case of the likelihood ratio statistic (LRT) should be used that examines the difference of goodness-of-fit between two log-linear models, i.e., the M_0 model of partial independence and the full M_1 model that represents a more general saturated alternative. The likelihood ratio statistics G^2 takes into account both models with a formula such as

$$G^2(M_0|M_1) = -2(L_0 - L_1)$$

where L_0 and L_1 are the likelihood functions (for more information, see [33]). G^2 is distributed as a chi-square distribution with degrees of freedom equal to the differences between both component models.

Statistical Analysis

We deemed the response strategy as more efficient with higher G^2 values achieved by subjects. We also fitted GLM to the behavioral data (monetary gains) to determine how an aware response strategy influenced winnings in the monetary and imaginary incentive cue conditions. The GLM fitting was performed with a *fitglm* function provided by the MATLAB statistics toolbox (Mathworks Inc.), and the ANOVA analysis was conducted with SPSS. For the GLM results, we reported t statistics and p values of individual effects in the model. The hypothesis testing procedures were complemented with effect size measures provided by the Matlab toolbox "Measures of Effect Size" by Hentschke and Stüttgen [15]. For computation of the effect size, we used Hedges' g parameters either for independent or dependent samples as recommended by Hentschke and Stüttgen [15].

Metacognitive Sensitivity of Wagering Based on "Type 2" Signal Detection Theory

To validate LLM outcomes, we also used an alternative measure of metacognitive awareness based on a "Type 2" signal detection theory: the so-called a *meta-d'* measure invented by Maniscalco and Lau [20, 21]. This type of sensitivity measure quantifies how much information is available for metacognition, and, in fact, can assess how efficient the confidence responses are at discriminating correct from incorrect classifications [20, 21]. It is emphasized that meta-d' can be thought of as a measurement of the signal which is available to perform a type 2 task such as discrimination accuracy. In the case of PDW usage, if observers' wagers are informative with regard to the correctness of first-order discriminations, the observer can be deemed as being metacognitively aware of the first-order representation. The calculations of meta-d' measures were conducted with the free Matlab code provided by Maniscalco and Lau [20]. It is important to mention that this measure of metacognition is based on the assumption that type 1 data is normally distributed, whereas the proposed LLM analysis uses a non-parametric assumption. Estimation of meta-d' was based on unequal variance SDT assumptions and, pertinently, the meta-d' parameter is a single measure of sensitivity that jointly corresponds to the areas under the response-specific type 2 ROCs [21].

Simulation of Wagering Behavior with Neural Network

To simulate choice wagering behavior, we employed a standard neural network modeling based on feed-forward architecture, as implemented in the Neural Network Toolbox provided in MATLAB. The feed-forward architecture consisted of one hidden layer of sigmoid neurons that was followed by an output layer of linear neurons (see Fig. 2). Such a network is typically used to approximate any function and its discontinuities [3]. In our case, the network served as a general approximator that was intended to learn the relations between four input wagers and the output of the corresponding LLM function. In the sense of a function approximation problem, the feed-forward network was trained to perform non-linear regression based on the mean square error used to estimate the error between the network output and the target output of the log-linear function G^2 (for more details, see [3]). The modeling was based on standard multi-layer networks with a non-linear hyperbolic tangent transfer function. The Neural Network toolbox was employed for matrix manipulation, network generation, training, as well as the evaluation stages. Before the process of network training, the data division was automatically performed, which is a standard practice in neural modeling [3]. In particular, the entire dataset was randomly divided into three disjunctive sets: the training set (80% of data samples), used for the neural network training; the

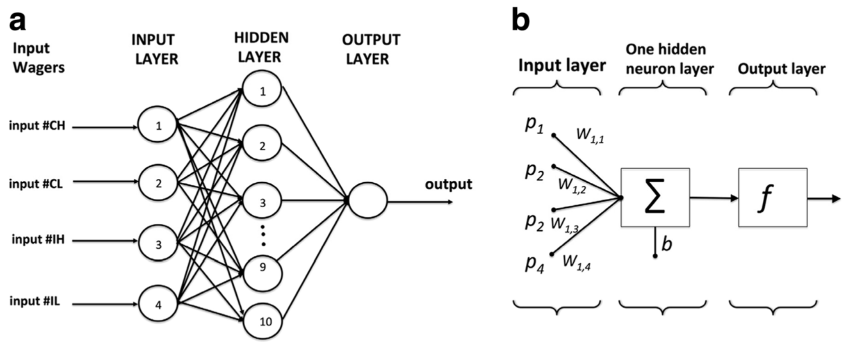

Fig. 2 Architecture of neural network unit. To model the data, we used a simple one-module architecture that captures the essence of the post-wagering task. The module takes wagers as an input (CH, CL, IH, IL) and produces log-linear statistics as output. **a** Initially, a network with 10 neurons in the hidden layer and a hyperbolic tangent transfer function was used. Standard Matlab-compatible tools for matrix manipulation, network generation, training, and evaluation were employed. As a result of training, the neural network along with architectural information and weight values was obtained. **b** The network architecture was optimized. The optimization results were surprising, since an extremely simple network was achieved that consisted of a single non-linear neuron with four inputs and one output

validation set (10% of data), used for current assessment of the network quality while training; and the test set, used for final testing of the neural network quality (10% of data).

The neural network simulation operated in two modes. In the learning mode, multiple samples of expected behavior, i.e., the contingency data represented by four wagers (inputs) along with its relevant G^2 value (outputs), were used to teach to the neural network. The learning mode was stopped when the assumed approximation error was reached or there was no further improvement in the quality of approximation. In addition, the learning proceeded by making noise adjustments by randomly removing inputs with a probability of 0.03. The variability of the input-output training sets was generated with a multinomial sampling procedure [1, 33]. In terms of the sampling procedure, variability of the contingency data was sampled with respect to the marginal proportions and the associations between variables. The cell probabilities were determined by fixing a total sample at 160 trials, and the input contingency data was sampled. To avoid the effect of overfitting the network to the data, it is advised to use as few neurons in the hidden layer as possible [3].

Initially, our modeling resulted in a neural network with ten neurons in the hidden layer (see Fig. 2a). Then, some optimization attempts were performed to reduce the number of hidden neurons and, as a result, a simple network architecture was achieved that consisted of one hidden neuron with four inputs and one output (see Fig. 2b). Then, the network modeling switched to "production mode" and the neural network along with its architectural information and weight values were complete. After a neural network simulation, the outcomes of input weights were extracted and submitted to further statistical analyses of GLM. For GLM fitting, to identify factors that best described wagering performance, we evaluated alternative combinations of the main effects and interactions between the main effects, and finally the best-supported model was chosen.

Results

Behavioral Data

First, we began our analysis by examining whether the incentive cue had a significant effect on wagering. The yes-no discriminations were reorganized into contingency tables accommodating both levels of wagers (high vs. low) and the accuracy (correct vs. incorrect response), accordingly. Then, we submitted the data to a mixed repeated-measures analysis of variance (ANOVA) and compared the responses from the wagering task between the two groups (i.e., monetary vs. imaginary incentives) across three within-subject factors such as stimulus durations (17 vs. 33 ms), accuracy (correct vs. incorrect), and wager (low vs. high). The analysis indicated no effect of the between-group factor (incentive cue) on performance, $F < 1$, and no interaction between this group factor and the stimulus duration $F(1,22) = 1.30$, $p > 0.05$, no interaction with the wager variable, $F(1,22) = 1.12$, $p > 0.05$, and no higher-order interaction with all within-subject factors, $F < 1$. We found the main effect for the accuracy of response, $F(1,22) = 71.42$, $p < 0.0001$, $\eta^2_p = 0.77$, the main effect of the wager, $F(1,22) = 8.87$, $p < 0.01$, $\eta^2_p = 0.29$, and the interaction between accuracy and the stimulus exposure, $F(1,22) = 49.46$, $p < 0.0001$, $\eta^2_p = 0.69$, the interaction between the time duration and wagering, $F(1,22) = 26.32$, $p < 0.0001$, $\eta^2_p = 0.55$, and the second-order interaction between the time exposure, accuracy, and wagering, $F(1,22) = 81.2$, $p < 0.05$, $\eta^2_p = 0.79$. These results suggested that there was no impact of the monetary incentive cue on wagering performance whatsoever.

Then, for each participant, we investigated whether response strategy was aware or unaware by examining partial associations between accuracy and wagering with the LRT statistics. For each stimulus duration, we computed the individual G^2 value and then assessed awareness of wagering strategies by checking whether the G^2 statistics reached a cut-off of 3.84 according to the chi-square distribution with 1 degree of freedom (p value of 0.05). For example, the G^2 statistics for subject S4 from the real monetary incentive group for the 17-ms exposure was 0.67. Because this value was deemed not significant, $p = 0.41$, this indicated that wagering strategy was unaware for this individual. In the case of the 33-ms condition, the hypothesis of partial association between accuracy and wagering was rejected for this participant, since the G^2 parameter of 4.38 yielded a p value of 0.04 that indicated that this participant used an aware strategy while wagering on the correctness of the first-order discriminations. To investigate whether wagering strategies were aware at the group level, we ran analysis of binomial data with a two-tail binomial test ($p < 0.05$), where the presence or absence of the aware strategy was coded as 1 or 0, respectively. The resulting two-tailed p value from this test indicated the chance of observing metacognitive awareness at the group level. In case of 17-ms targets, 14 participants (14/24) showed awareness, which was a chance of $p = 0.54$, revealing the fact that strategies were unaware when wagered on the correctness of their decisions at the group level. For the 33-ms presentation, we observed that 21 participants (21/24) performed successfully in the post-decision wagering task ($p < 0.001$, two-tail binomial test). Taking these results together, we concluded that for the 17-ms targets, participants at the group level tended to use unaware wagering strategies when judging the correctness of their decisions, while for the 33-ms targets, participants tended to use aware wagering strategies. Finally, we also examined how wagering strategies intended to maximize monetary profits were dependent on the awareness factor. The results of the GLM fitting are summarized in Table 1. There was evidence that there was an important increase in monetary gains for the predictor variable of awareness, $t(42) = 2.49$, $p < 0.05$, and no other main effect or interaction was significant. This result indicated that aware strategies substantially contributed to monetary gains in the post-wagering task.

Table 1 Analysis of behavioral data with generalized linear regression model

Generalized linear regression model:
Monetary_gain ~ 1 + awareness * incentive_cue
Distribution = normal

Estimated coefficients:

	Estimate	SE	t stat	p value
(Intercept)	−1.9147	5.0185	−0.38153	0.70473
Awareness	6.8875	2.76	2.4955	0.016594
Incentive_cue	0.13528	3.0948	0.043711	0.96534
Awareness:Incentive_cue	−0.24847	1.7071	−0.14555	0.88497

46 observations, 42 error degrees of freedom. Estimated dispersion, 6.06. F statistic vs. constant model, 19.6; p value = 4.2e−08

Comparison Between Metacognitive Sensitivity and LLM Partial Associations

In the next step of our analysis, we validated the LLM results with the alternative measure of awareness by applying an analysis of meta-d' sensitivity to wagering behavior. Initially, we investigated metacognition by submitting the meta-d' values into a mixed repeated ANOVA that indicated no effect of the incentive on metacognition, $F < 1$. Furthermore, we examined simple effects for both stimulus durations. For the 17-ms targets, we found that metacognitive sensitivity for the monetary group yielded a mean meta-d' of 0.57 (SD = 0.47), whereas for the imaginary group the meta-d' sensitivity was 0.49 (SD = 0.47). As opposed to the shorter stimulus exposure, metacognitive capacities increased significantly for the 33-ms presentation, resulting in larger values of meta-d' for either the monetary group, 1.48 (SD = 1.03) (paired t test, $p < 0.05$, Hedges' $g = 0.85$), or the imaginary group, 1.40 (SD = 0.75) (paired t test, $p < 0.001$, Hedges' $g = 1.40$). We then examined metacognitive awareness at the group level by checking whether the zero metacognitive sensitivity was included in the 95% confidence interval. According to one-sample t test statistics, the monetary group for the 17-ms targets exhibited a lack of metacognitive awareness, $t(11) = 1.98$, $p > 0.05$, $g_1 = 0.57$, as opposed to the 33-ms targets, for which metacognitive awareness was present, $t(11) = 4.94$, $p < 0.001$, $g_1 = 1.43$. For the imaginary group, the results indicated that participants used aware wagering strategies either for 17-ms targets, $t(11) = 3.63$, $p < 0.01$, $g_1 = 1.05$, or for 33-ms conditions, $t(11) = 6.47$, $p < 0.0001$, $g_1 = 1.05$. This suggested that the SDT and LLM measures were consistent with indicating awareness for the 33-ms targets, although according to the SDT measure the imaginary incentive group was aware for 17-ms as opposed to LLM. Furthermore, we plotted both metacognitive measures side by side (see Fig. 3a, b). Strikingly, as can be seen, both metacognitive measures for both stimulus exposures show very similar patterns across all subjects. Furthermore, examination of the correlation coefficients between partial association outcomes (G^2 values) and meta-d' sensitivities indicated that the SDT and LLM measures were highly correlated. In the case of the monetary group, the correlation coefficients yielded values of $r = 0.77$, $p < 0.01$, and $r = 0.72$, $p < 0.01$ for 17- and 33-ms conditions, respectively. In the case of the imaginary group, the correlation coefficients reached $r = 0.74$, $p < 0.01$, and $r = 0.92$, $p < 0.0001$ for 17- and 33-ms conditions, respectively. This suggested that the predictions given by SDT and LLM analyses were largely overlapping.

Neural Simulation of Wagering Behavior

For each individual, we then analyzed wagering behavior at the 17- and 33-ms conditions by inspecting input weights of

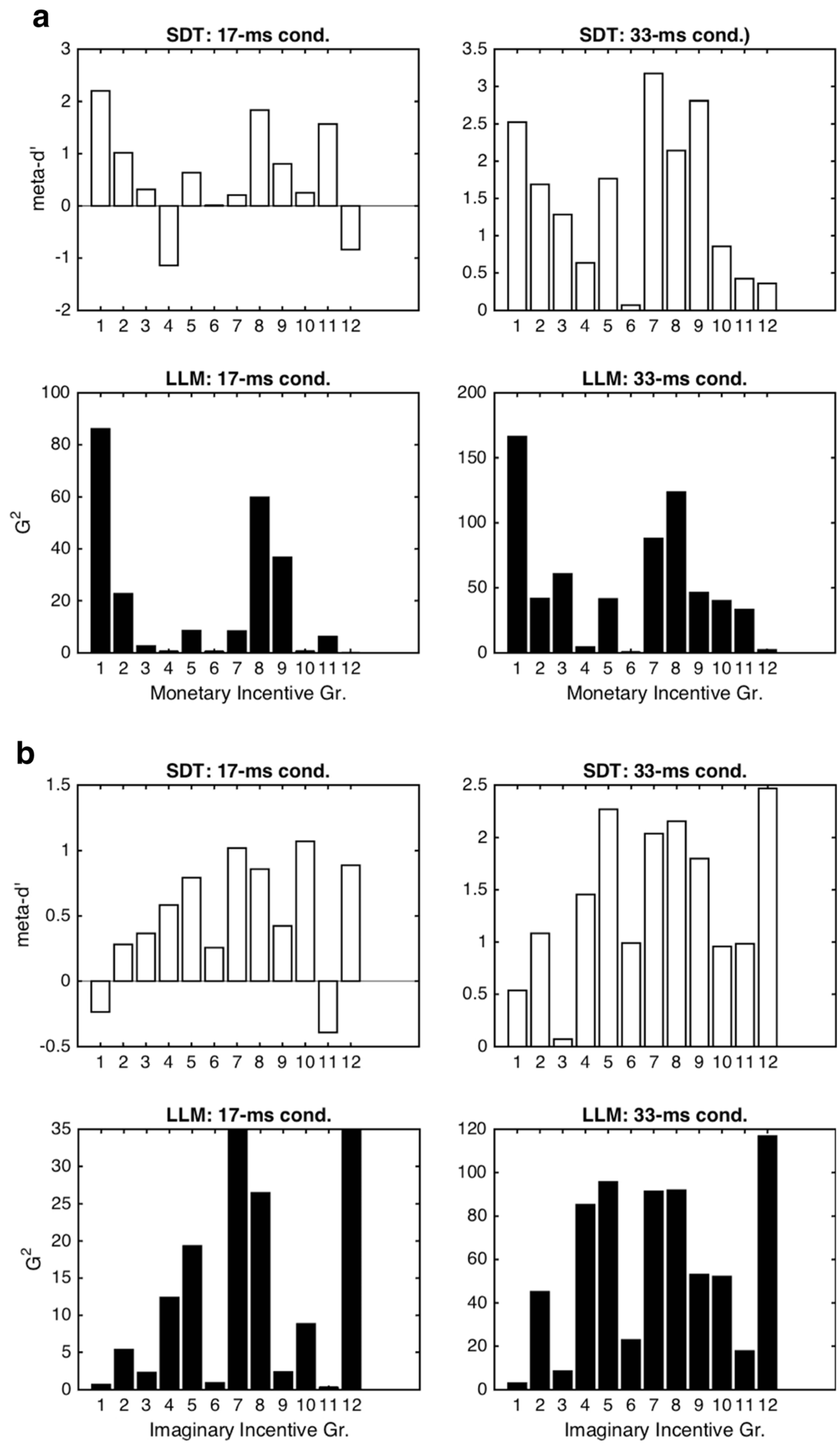

Fig. 3 **a** Comparison between metacognitive sensitivity and LLM partial associations for the "real incentive" group. The validation of the LLM results with meta-d' sensitivity indicated that both measures largely overlapped. **b** Comparison between metacognitive sensitivity and LLM partial associations for the "imaginary incentive" group

the neural network unit. The resulting input weights across all subjects are presented in Fig. 4. As can be seen, after averaging, distinguished patterns of initial weights appeared for correct and incorrect low wagers. Note that the other two input weights received negligible values and were not considered for further investigations. In particular, for the monetary incentive group, for the correct low wagers the values of input weights were 0.57 and 0.26 for the 17- and 33-ms conditions, respectively; for the incorrect low wagers, the input weights were −0.81 and −0.52, for the 17- and 33-ms conditions, respectively. For the imaginary incentive group, we found that the neural input weights were 0.34 and 1.08 for correct low wagers, whereas for incorrect responses the input weights were equal to −0.58 and −1.14 for 17- and 33-ms conditions, respectively. This indicated that the input weights for correct low wagers could be associated with excitation, while the

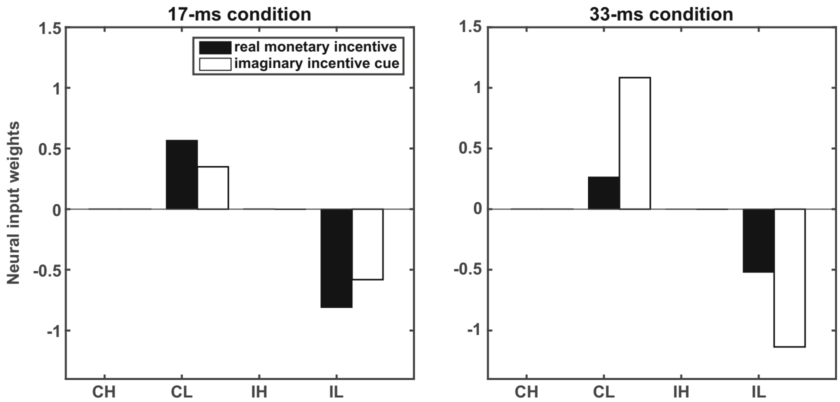

Fig. 4 Results of simulation for wagering performance. Neural input weights for 17- and 33-ms conditions. *CH* correct response and high wager, *CL* correct response and low wager, *IH* incorrect response and high wager, *IL* incorrect response and low wager. The simulation results indicated that correct low wagers are characterized by positive value weights, which indicates some kind of excitation, while for incorrect values, the neural input weights showed patterns of excitation

input of the neural network for handling incorrect low wagers was associated with inhibition behavior. To investigate further how activation was encoded in the neural network module across the stimulus durations, we inspected variability of simulated inhibitory and excitatory input weights with the Mann-Whitney *U* tests ($p < 0.05$). No differences were observed for the excitatory and inhibitory neural inputs for each group at each stimulus exposure. This suggested similar patterns of neural network activation when engaging unaware and aware strategies.

Finally, we investigated the impact of the neural network on wagering behavior with the GLM model (see Table 2). Here, the dependent variable was the amount of gambled money predicted by the positive and negative input weights of the network and the stimulus duration; the incentive cue categorical variable was excluded from the model to achieve the best-supported GLM model. The analysis revealed increases of monetary gains for the excitatory neural input, $t(41) = 5.11$, $p < 0.0001$, for the inhibitory neural input, $t(41) = 4.87$, $p < 0.0001$ as well as for the stimulus duration, $t(41) = 3.11$, $p < 0.01$. There was also evidence of decreases in winnings for the interaction of the excitatory input and the time stimulus, $t(41) = -3.32$, $p < 0.01$, for the interaction of inhibitory input and the stimulus duration, $t(41) = -3.20$, $p < 0.01$. Note that the GLM model included categorical variables of stimulus durations; the 17-ms condition was used as the reference level, whereas the dummy variable was employed for the 33-ms condition. Hence, GLM prediction of the decrease for this second interaction term indicated that the inhibitory weight parameter (negative value) contributed to increasing monetary gains for the 33-ms condition. The GLM also predicted that there was an increase for the second-order interaction between both input weights and the stimulus exposure, $t(41) = 4.37$, $p < 0.01$, and again this term contributed to lowering monetary gains due to the negative value of inhibitory weight. These interactions with longer exposures suggested that the neural network unit allowed more effective strategies by letting participants increase their monetary gains.

Discussion

The contemporary connectionist analysis for empirical studies on human faces focuses either on isolating perceptual and emotional processes engaged in detection of facial expressions (see e.g., [22]) or proposes more complex frameworks that aim to capture a variety of different aspects related to emotion processing (e.g., action, motivation and metacognition—see e.g., [32]). Nevertheless, the present research concentrated on the metacognitive awareness of perceiving facial expressions, which is a topic rarely analyzed with connectionist approaches. In particular, in this study, we employed a new method of measuring awareness based on log-linear analysis to test the association between wagering and accuracy in detecting masked facial expressions. Following the connectionist approach, we further used the LLM results to simulate activation of the neural network for the higher-order classifications that underlie aware and unaware wagering behavior. The resulting neural network allowed us to predict wagering outcomes based on log-linear accuracy. Surprisingly, the neural network simulations were effectively performed with a simple architecture consisting of one hidden neuron unit with input weights sensitive to low wagers.

Our behavioral results show that participants were more inclined to employ effective response strategies to increase winnings when they were aware of their processing results, i.e., first-order discrimination outcomes for the 33-ms facial targets. In addition, these patterns of responding were supported by the neural network performance that underlies wagering responses. Indeed, it can be seen from the GLM analysis that participants made more monetary gains via wagering when

Table 2 Analysis of behavioral data with generalized linear regression

Gain ~ [linear formula with 7 terms in 3 predictors]
Distribution = normal
Estimated coefficients:

	Estimate	SE	*t* stat	*p* value
(Intercept)	145.21	19.753	7.3511	5.2882e−09
Excitatory_weight	305.02	59.634	5.1149	7.7448e−06
Inhibitory_weight	315.32	64.781	4.8675	1.7161e−05
Stimulus_duration	81.646	26.26	3.1091	0.0034059
Excitatory_weight:Stimulus_duration	−253.52	76.356	−3.3202	0.0018963
Inhibitory_weight:Stimulus_duration	−263.67	82.521	−3.1951	0.0026886
Excitatory_weight:Inhibitory_weight:Stimulus_duration	51.229	11.727	4.3686	8.3141e−05

48 observations, 41 error degrees of freedom. Estimated dispersion, 1.47e+03. *F* statistic vs. constant model, 21.5; *p* value = 3.17e−11

the longer stimulus duration was applied. Clearly, our results suggest that larger input weights of the network (we observed three significant interaction terms in the model) must arrive at the network unit to promote aware response. Taken together, both behavioral and simulation results can lead us to conclude that awareness enables more effective post-decision wagering response strategies by providing participants with the ability to increase their monetary gains.

In this work, in contrast to Pasquali et al. [24], we focus on evaluation of individual performance by adapting the neural network only to metacognitive judgments. Thus, we develop the connectionist model of metacognitive ratings on visual identification that aimed to test how participants rate awareness with post-decision wagering, but not the sensory perception of emotional stimuli. The important property of our neural network model was its feed-forward architecture with one hidden neuron layer. Here, the inputs were wagers' sequences presented to the neural network, while the target outputs were represented by a log-linear function and its relevant G^2 values. The input-output relationships were then subjected to a feed-forward learning algorithm that adapted the input weights of the network to approximate the wagering behavior. Thus, the resulting input weights of the network were preconditions for wagers on facial discrimination to reach consciousness.

Interestingly, our neural network simulation results clearly show that only input weights of the network associated with low wagers significantly contributed to wagering performance in both the aware or unaware conditions. In other words, only low wager inputs were informative in predicting metacognitive awareness. Interestingly, for the subset of the low wager inputs, correct responses were associated with positive input weights, while incorrect responses were associated with negative weights. Thus, at the neural network level, wagering performance is informed mainly by the correctness of low wagers. The high wagers do not discriminate between aware and unaware states, since they are observed mainly in conscious conditions. Such results may speak against applying post-decision wagering to studies investigating awareness. This arises from the fact that wagering is predictive for awareness only at low wagers so that low wagers may be associated with both aware and unaware conditions. This is additional evidence that post-decision wagering could be potentially influenced by loss aversion. Thus, if post-decision wagering studies are to be used in consciousness studies further investigations should instead be focused on scrutiny of low wagers, as suggested elsewhere (see also [7]).

We have not observed any difference in conscious response strategies depending on the type of incentives applied. Wagering behavior with imaginary and real money was comparable. This confirms the claim by Persaud et al. [26], who argued that wagering with imaginary money is as effective as betting with real money. Similarly, the connectionist model did not confirm that real money incentives are more influenced by response strategy, as it was suggested by Dienes and Seth [11]. Thus, one is free to choose which version of the method should be applied in a particular study. However, if real incentives are needed in the experimental design, our study suggests that outcomes should not to be influenced by loss aversion more than under imaginary incentive conditions. Thus, our version of the PDW method may be successively applied in studies involving participants with limited reporting abilities, particularly for whom the real incentives may be easier to apply over the ongoing task procedure.

To sum up, our LLM analysis and neural network simulation allowed us to confirm and further conduct in-depth investigations of loss aversion, which is commonly observed in post-decision wagering studies (see e.g., [7, 31]). It seems that in both cases—either for real or imaginary monetary incentives—it is likely that participants start to believe that gambling will lead to a loss. Hence, post-decision with imaginary money may be reasonably effective in experiments designed to examine awareness. Thus, in our study, we propose a connectionist model of post-decision wagering that is not ultimately intended to simulate meta-knowledge establishing awareness but rather focuses on the impact of loss aversion on aware and unaware wagering. Our work also demonstrates that neural networks may be applicable, not only to investigate low-level phenomena related to emotion (see [32]) but also to reveal a computational view on patterns of association between awareness and advantageous wagering.

Acknowledgements This work was supported by the National Science Center under grants' decision DEC-2011/03/B/HS6/01799 to R.S. We thank Tomasz Guszkowski for designing neural computation.

Compliance with Ethical Standards

Funding This study was funded by the National Science Center (grant number 2011/03/B/HS6/01799).

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

1. Agresti A. An introduction to categorical data analysis (Vol. 423). Wiley-Blackwell. 2007.
2. Anderson G, Brown RIF. Real and laboratory gambling, sensation-seeking and arousal. Br J Psychol. 1984;75(3):401–10.
3. Beale, M. H., Hagan, M. T., & Demuth, H. B. (2015). Neural Network Toolbox™ getting started guide.
4. Cleeremans A. Consciousness: the radical plasticity thesis. Prog Brain Res. 2008;168:19–33.
5. Cleeremans A, Haynes JD. Correlating consciousness: a view from empirical science. Rev Int Philos. 1999:387–420.
6. Cleeremans A, Timmermans B, Pasquali A. Consciousness and metarepresentation: a computational sketch. Neural Networks: the official journal of the International Neural Network Society. 2007;20(9):1032–9. doi:10.1016/j.neunet.2007.09.01.
7. Clifford CW, Arabzadeh E, Harris JA. Getting technical about awareness. Trends Cogn Sci. 2008;12(2):54–8. doi:10.1016/j.tics.2007.11.00.
8. Dailey MN, Cottrell GW, Padgett C, Adolphs R. EMPATH: a neural network that categorizes facial expressions. J Cogn Neurosci. 2002;14(8):1158–73.
9. Dehaene S, Changeux JP. Experimental and theoretical approaches to conscious processing. Neuron. 2011;70(2):200–27.
10. Dienes Z, Perner J. Assumptions of a subjective measure of consciousness: three mappings. In: Gennaro R, editor. Higher order theories of consciousness. Amsterdam: John Benjamins Publishers; 2004. p. 173–99.
11. Dienes Z, Seth A. Gambling on the unconscious: a comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. Conscious Cogn. 2010;19(2):674–81.
12. Ekman P, Friesen WV. Pictures of facial affect. Palo Alto, CA: Consulting Psychologists Press; 1976.
13. Fleming SM, Dolan RJ. Effects of loss aversion on post-decision wagering: implications for measures of awareness. Conscious Cogn. 2010;19(1):352–63.
14. Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G. Relating introspective accuracy to individual differences in brain structure. Science (New York, NY). 2010;329(5998):1541–3.
15. Hentschke H, Stüttgen MC. Computation of measures of effect size for neuroscience data sets. Eur J Neurosci. 2011;34(12):1887–94.
16. Ishai A, Pessoa L, Bikle PC, Ungerleider LG. Repetition suppression of faces is modulated by emotion. Proc Natl Acad Sci U S A. 2004;101:9827–32.
17. Lisetti C L, Rumelhart DE. Facial expression recognition using a neural network. In FLAIRS Conference; 1998. 328–332.
18. Lundqvist D, Öhman A. Emotion regulates attention: the relation between facial configurations, facial emotion, and visual attention. Vis Cogn. 2005;12(1):51–84.
19. Lundqvist D, Flykt A, Öhman A. The Karolinska Directed Emotional Faces (KDEF). CD ROM from Department of Clinical Neurosci Psychol Sec Karolinska Inst; 1998. 91–630
20. Maniscalco B, Lau H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Conscious Cogn. 2012;21(1):422–30. doi:10.1016/j.concog.2011.09.021.
21. Maniscalco B, Lau H. Signal detection theory analysis of type 1 and type 2 data: meta-d', response-specific meta-d', and the unequal variance SDT mode. In S. M. Fleming & C. D. Frith (Eds.), The Cognitive Neuroscience of Metacognition; 2014. 25–66. Springer.
22. Mermillod M, Vermeulen N, Lundqvist D, Niedenthal PM. Neural computation as a tool to differentiate perceptual from emotional processes: the case of anger superiority effect. Cognition. 2009;110(3):346–57.
23. Öhman A. The role of the amygdala in human fear: automatic detection of threat. Psychoneurendocrinology. 2005;30:952–8.
24. Pasquali A, Timmermans B, Cleeremans A. Know thyself: metacognitive networks and measures of consciousness. Cognition. 2010;117(2):182–90. doi:10.1016/j.cognition.2010.08.01.
25. Peters MA, Lau H. Human observers have optimal introspective access to perceptual processes even for visually masked stimuli. ELife; 2015. 4. doi:10.7554/eLife.0965
26. Persaud N, McLeod P, Cowey A. Post-decision wagering objectively measures awareness. Nat Neurosci. 2007;10(2):257–61.
27. Pessoa L. To what extent are emotional visual stimuli processed without attention and awareness? Curr Opin Neurobiol. 2005;15(2):188–96.
28. Ruffman T, Garnham W, Import A, Connolly D. Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. J Exp Child Psychol. 2001;80(3):201–24. doi:10.1006/jecp.2001.263.
29. Sandberg K, Bibby BM, Timmermans B, Cleeremans A, Overgaard M. Measuring consciousness: task accuracy and awareness as sigmoid functions of stimulus duration. Conscious Cogn. 2011;20(4):1659–75. doi:10.1016/j.concog.2011.09.002.
30. Sato YD, Nagatomi T, Horio K, Miyamoto H. The cognitive mechanisms of multi-scale perception for the recognition of extremely similar faces. Cogn Comput. 2015;7(5):501–8.
31. Schürger A, Sher S. Awareness, loss aversion, and post-decision wagering. Trends Cogn Sci. 2008;12(6):209–10. doi:10.1016/j.tics.2008.02.01.
32. Sun R, Wilson N, Lynch M. Emotion: a unified mechanistic interpretation from a cognitive architecture. Cogn Comput. 2016;8(1):1–14.
33. Szczepanowski R. Absence of advantageous wagering does not mean that awareness is fully abolished. Conscious Cogn. 2010;19(1):426–31.
34. Szczepanowski R, Pessoa L. Fear perception: can objective and subjective awareness measures be dissociated? J Vis. 2007;7(4):1–17.
35. Szczepanowski R, Traczyk J, Wierzchoń M, Cleeremans A. The perception of visual emotion: comparing different measures of awareness. Conscious Cogn. 2013;22(1):212–20. doi:10.1016/j.concog.2012.12.00.

36. Timmermans B, Schilbach L, Pasquali A, Cleeremans A. Higher order thoughts in action: consciousness as an unconscious re-description process. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 2012;367(1594):1412–23. doi:10.1098/rstb.2011.042.
37. Wierzchoń M, Asanowicz D, Paulewicz B, Cleeremans A. Subjective measures of consciousness in artificial grammar learning task. Conscious Cogn. 2012;21(3):1141–53. doi:10.1016/j.concog.2012.05.01.
38. Wierzchoń M, Wronka E, Paulewicz B, Szczepanowski R. Post-decision wagering affects metacognitive awareness of emotional stimuli: an event related potential study. PLoS One. 2016;11(8): e0159516. doi:10.1371/journal.pone.015951.
39. Zotto MD, Pegna AJ. Processing of masked and unmasked emotional faces under different attentional conditions: an electrophysiological investigation. Front Psychol. 2015;6:1691. doi:10.3389/fpsyg.2015.0169.