

SCIENTIFIC REPORTS



OPEN

Post-response $\beta\gamma$ power predicts the degree of choice-based learning in internally guided decision-making

Takashi Nakao¹, Noriaki Kanayama^{2,3}, Kentaro Katahira⁴, Misaki Odani⁵, Yosuke Ito⁵, Yuki Hirata⁵, Reika Nasuno⁵, Hanako Ozaki⁵, Ryosuke Hiramoto⁵, Makoto Miyatani¹ & Georg Northoff⁶

Received: 29 April 2016
Accepted: 09 August 2016
Published: 31 August 2016

Choosing an option increases a person's preference for that option. This phenomenon, called choice-based learning (CBL), has been investigated separately in the contexts of internally guided decision-making (IDM, e.g., preference judgment), for which no objectively correct answer exists, and externally guided decision making (EDM, e.g., perceptual decision making), for which one objectively correct answer exists. For the present study, we compared decision making of these two types to examine differences of underlying neural processes of CBL. As IDM and EDM tasks, occupation preference judgment and salary judgment were used, respectively. To compare CBL for the two types of decision making, we developed a novel measurement of CBL: decision consistency. When CBL occurs, decision consistency is higher in the last-half trials than in first-half trials. Electroencephalography (EEG) data have demonstrated that the change of decision consistency is positively correlated with the fronto-central beta–gamma power after response in the first-half trials for IDM, but not for EDM. Those results demonstrate for the first time the difference of CBL between IDM and EDM. The fronto-central beta–gamma power is expected to reflect a key process of CBL, specifically for IDM.

How value is learned through the decision-making sequence is an important question that is explored broadly in psychology and neuroscience^{1–6}. Many studies of decision making have used tasks in which a particular more-or-less predictable answer is available^{7–10}. These are called externally guided decision-making tasks (EDM)^{11,12}. In such situations, people must adjust their decision criteria to choose an externally defined single correct answer. Studies of decision making of this type have revealed that learning processes of item value have generally been explained by reinforcement learning theory, whereby the expected value (i.e., the magnitude of outcome times the probability of the outcome) guides the decision. The expected value is updated by the prediction error (i.e., discrepancies between expected and actual outcome)^{13–15}. Recently, reports of studies using electroencephalography (EEG) have described that increased fronto-central beta–gamma power is visible after positive feedback in reward-based decision-making tasks. That increased power is thought to reflect positive reinforcement processes^{16–24}.

In addition to EDM, other types of decision making in which no objective correct answer is available^{25–31}. They constitute internally guided decision making (IDM)¹¹. Such decisions are usually made in the context of preference judgment^{32–35} as well as in the context of moral judgment^{36–38}, where the answer depends on the subject's own, i.e., *internal*, preferences rather than on circumstantial, i.e., *external*, criteria. Although no prediction error is available in the IDM, the value or preference of an item is known to be altered through the results of a person's own decision. The preference for the chosen item is increased, but that for the rejected item is decreased^{39–43}. This phenomenon is designated in the literature as choice-induced preference change^{39,41,44,45}.

Recently, as with the EDM, a reinforcement-learning-like mechanism is thought to explain how choice-induced preference change occurs^{46–48}. Akaishi *et al.*⁴⁶ reported that people tend to repeat the same decision when required to make a decision based on perceptually ambiguous stimulus with no feedback.

¹Department of Psychology, Graduate School of Education, Hiroshima University, Hiroshima, Japan. ²Department of Psychiatry and Neurosciences, Institute of Biomedical & Health Sciences, Hiroshima University, Hiroshima, Japan. ³Center of KANSEI Innovation, Hiroshima University, Hiroshima, Japan. ⁴Department of Psychology, Graduate School of Environment, Nagoya University, Nagoya, Aichi, Japan. ⁵Faculty of Education, Hiroshima University, Hiroshima, Japan. ⁶Institute of Mental Health Research, University of Ottawa, Ottawa, Canada. Correspondence and requests for materials should be addressed to T.N. (email: tnakao@hiroshima-u.ac.jp)

Moreover, Akaishi *et al.*⁴⁶ demonstrated that the underlying mechanism of that tendency is explained well by a reinforcement-learning-based model in which the likelihood of choosing an option is updated by the choice itself, called choice-based learning (CBL). Although the perceptual decision-making task used by Akaishi *et al.*⁴⁶ was more of EDM than IDM, they proposed that the same mechanism can explain the choice-induced preference change: in a situation where no externally defined single correct answer exists, choosing a particular item is followed by reinforcement of the chosen item as if the own choice is regarded as a correct answer. This process engenders an increase in the preference for the chosen item and an increase in the likelihood of choosing the same item at the following opportunity. Consistent with the proposal by Akaishi *et al.*⁴⁶, neuroimaging studies of choice-induced preference change have shown activation within typical neural substrates of reward-based reinforcement learning (e.g., ventral striatum, dorsal anterior cingulate cortex, and ventromedial prefrontal cortex)^{40,49–52}. However, those earlier studies examined brain activities after stimulus onset for free choice between two items^{50,52} or for preference ratings of respective items^{40,49,51} using functional magnetic resonance imaging (fMRI). Consequently, although neural activities related to choice response are expected to play a crucially important role in CBL because the choice itself induces the value change, the relation between CBL and neural activities related to the response remains unclear.

Unfortunately, IDM and EDM have been investigated using different paradigms in terms of stimuli, tasks¹¹, and means to measure the change of preference or value through decision making. In IDM studies, choice-induced preference change has been investigated using the free choice paradigm⁴³, which includes pre-rating, free choice, and post-rating phases. In the pre-rating phase, the participant rates several items according to the participant's own preferences. In the free choice phase, two items are presented. Participants are asked to choose the preferred one. In the post-rating phase, a participant is asked to rate all items again in the same way as with pre-ratings. In this paradigm, decision making (i.e., free choice) for a particular item is required once. Then the value change is measured by subjective rating^{39,41,45}. In contrast, in EDM studies, decision making for the same stimuli is repeated several times to ascertain the choice-induced change in the likelihood of choosing the stimuli at the subsequent opportunity^{7,8,46,53}. Those differences have made it difficult to compare decision making of these two types. Consequently, no report in the relevant literature has described comparison of CBL between these two types of decision making within the same paradigm. For that reason, the differences between the decision making types in terms of the underlying neural processes of CBL remain to be elucidated.

This study was undertaken to compare IDM to EDM in terms of neural activities around choice response related to CBL. To compare CBL of these two types of decision making, we developed a novel measurement, i.e., the change of decision consistency, which is applicable to both IDM and EDM tasks (for more details, see Fig. 1(c) and *Methods*). The decision consistency is expected to be higher in the last half of the sequence of decision making than in the first half if the likelihood of choosing a chosen item is increased, or if the likelihood of choosing a rejected item is decreased. By conducting simulation using a simple CBL model, we confirmed that an increase of decision consistency was observed when CBL occurs (for details, see Supplemental Information and Fig. S1(c–e) for Simulation 2). Although the free choice paradigm is known to have a methodological flaw by which choice-induced preference change is observed even if no preference change exists^{42,44,54}, the change of decision consistency can avoid that problem. We also confirmed this point by conducting a simulation (Supplemental Information and Fig. S1(a,b) for Simulation 1). Actually, we expect that increased decision consistency is observed both in IDM and in EDM because CBL has been observed in IDM (i.e., choice-induced preference change)^{39,41,45} and EDM⁴⁶.

Additionally, we expect that a difference of CBL between IDM and EDM is observed in the relation between the change of decision consistency and the neural activities related to the response. Although no strong activation within the reward-related neural substrates was reported in CBL of EDM⁴⁶, it was reported in CBL of IDM^{40,49–52}. This evidence suggests that CBL in IDM is associated with reinforcement processes such as reward-based reinforcement learning, different from CBL in EDM, even if the item is not visibly reinforced by the externally delivered reward. For measurement of neural processes related to choice response, we specifically examined the beta–gamma (14–60 Hz) power for the response. As described above, earlier studies of reward-based decision making showed increased beta–gamma power to around 200–500 ms after presenting positive feedback. That range is thought to reflect positive reinforcement processes^{16–24}. Participants who showed greater beta–gamma power around 200–500 ms after response in the first-half trials are expected to show more CBL effects (i.e., increased decision consistency in the last-half trials than in first-half trials) in IDM, but not in EDM.

Results

Validation of the occupation preference judgment task as IDM, and salary judgment task as EDM.

As IDM and EDM tasks, we respectively used occupation preference judgment^{12,34} (e.g., “Which occupation would you rather do? ... Lawyer or Designer”), for which no objectively correct answer exists, and salary judgment (e.g., “Which occupation is highly paid? ... Lawyer or Designer”) for which one objectively correct answer exists (Fig. 1(a)). To confirm the decision criteria for each task, we computed the pre-rating–decision consistency (i.e., how often participants' decisions of preference judgment were consistent with pre-ratings of the subjective occupational preference; see Supplemental Information for detailed methods), and average annual salary database–decision consistency (i.e., how often participants' decisions were consistent with the actual average annual salary, which is based on a statistical survey by the Ministry of Health, Labour and Welfare of Japan; see Supplemental Information for details). As Fig. 1(b) shows, results of these indexes confirmed that the occupation preference judgment was based on internal criteria (i.e., subjective occupation preference) different from the salary judgment, whereas the salary judgment was based on the external criteria (i.e., the average salary of the presented occupation) different from the preference judgement. Two-way repeated-measures ANOVA (two decision tasks (IDM, EDM) × two criteria (subjective occupation preference, the average salary of the presented occupation)) revealed a significant interaction ($F(1, 23) = 124.85; p < 0.001, \eta_g^2 = 0.64$). Post-hoc tests of

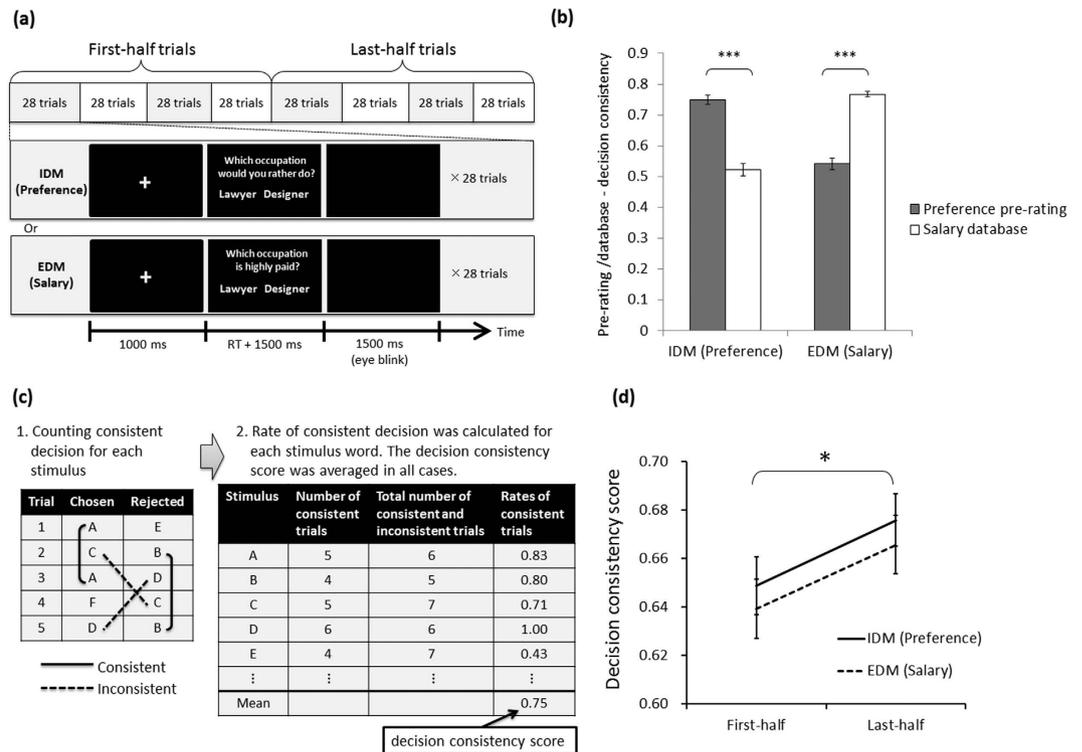


Figure 1. (a) Design of experimental tasks. As the internally guided decision making (IDM) and externally guided decision making (EDM) tasks, we used occupation preference judgment^{12,34}, for which no objectively correct answer exists, and salary judgment for which one objectively correct answer exists, respectively. Participants performed counterbalanced tasks of two types. RT denotes the reaction time. (b) Pre-rating–decision consistency (i.e., how often participants’ decisions of preference judgment were consistent with pre-ratings of the occupational preference), and average annual salary database–decision consistency (i.e., how often participants’ decisions were consistent with the actual average annual salary, which is based on a statistical survey by the Ministry of Health, Labour and Welfare of Japan) for each decision-making task (see *Supplemental Information* for more detailed methods to calculate these indexes). ***Denotes a significant main effect of epoch ($p < 0.0001$). Error bars show standard errors. (c) Schematic figure of the calculation of decision consistency score. The decision consistency score represents the rate of trials in which a certain occupation word was chosen or rejected repeatedly. In cases where a participant chose A (first trial in the example of this figure) and it was chosen again in the trial in which A was presented the next time (third trial in this example), we counted that trial as a consistent decision. In addition, in cases where a participant rejected B (second trial in this example) and it was rejected again in the trial in which B was presented the next time (fifth trial in this example), we counted that trial as a consistent decision. The consistent decisions were counted for each occupation word. Then that number was converted to a rate of consistent decision by dividing the number of consistent decisions by the sum of the number of consistent and inconsistent decisions. The average of the rate of consistent decisions was calculated across all occupation words. (d) Decision consistency scores for the first-half and the last-half trials in IDM (preference judgment) and EDM (salary judgment) tasks. *Denotes a significant main effect of epoch ($p < 0.01$). Error bars show standard errors.

the interaction revealed that the pre-rating–decision consistency was higher in the IDM (preference judgment) task than in the EDM (salary judgment) tasks ($p < 0.001$) in the dimensions of subjective occupation preference. For the dimension of average annual salary, the average annual salary database–decision consistency was higher in the EDM (salary judgment) task than in the IDM (preference judgment) task ($p < 0.001$). These results indicate that participants used internal or external criteria, respectively, for the preference judgment and salary judgment.

Decision consistency. The decision consistency score represents the ratio of trials in which a certain occupation word was repeatedly chosen or rejected (for details see Fig. 1(c) and *Methods*). Figure 1(d) presents a summary of the mean decision consistency score. It shows that decision consistency was higher in the respective last-half trials than in the first-half trials both in IDM (preference judgment) and EDM (salary judgment) tasks. Consistent with this observation, two-way repeated-measures ANOVA (two decision tasks (IDM, EDM) \times two epochs (first-half and last-half trials)) revealed a significant main effect of epoch ($F(1, 23) = 8.20, p = 0.009, \eta_G^2 = 0.05$). No significant main effect of decision task ($F(1, 23) = 0.50, p = 0.49, \eta_G^2 = 0.01$) or interaction ($F(1, 23) = 0.001, p = 0.97, \eta_G^2 = 0.00$) was found.

Regarding the results of the trial-based decision consistency score, reaction times, and pre-rating data including subjective rating for average annual salary, see *Supplemental Information* (Supplementary Fig. S3).

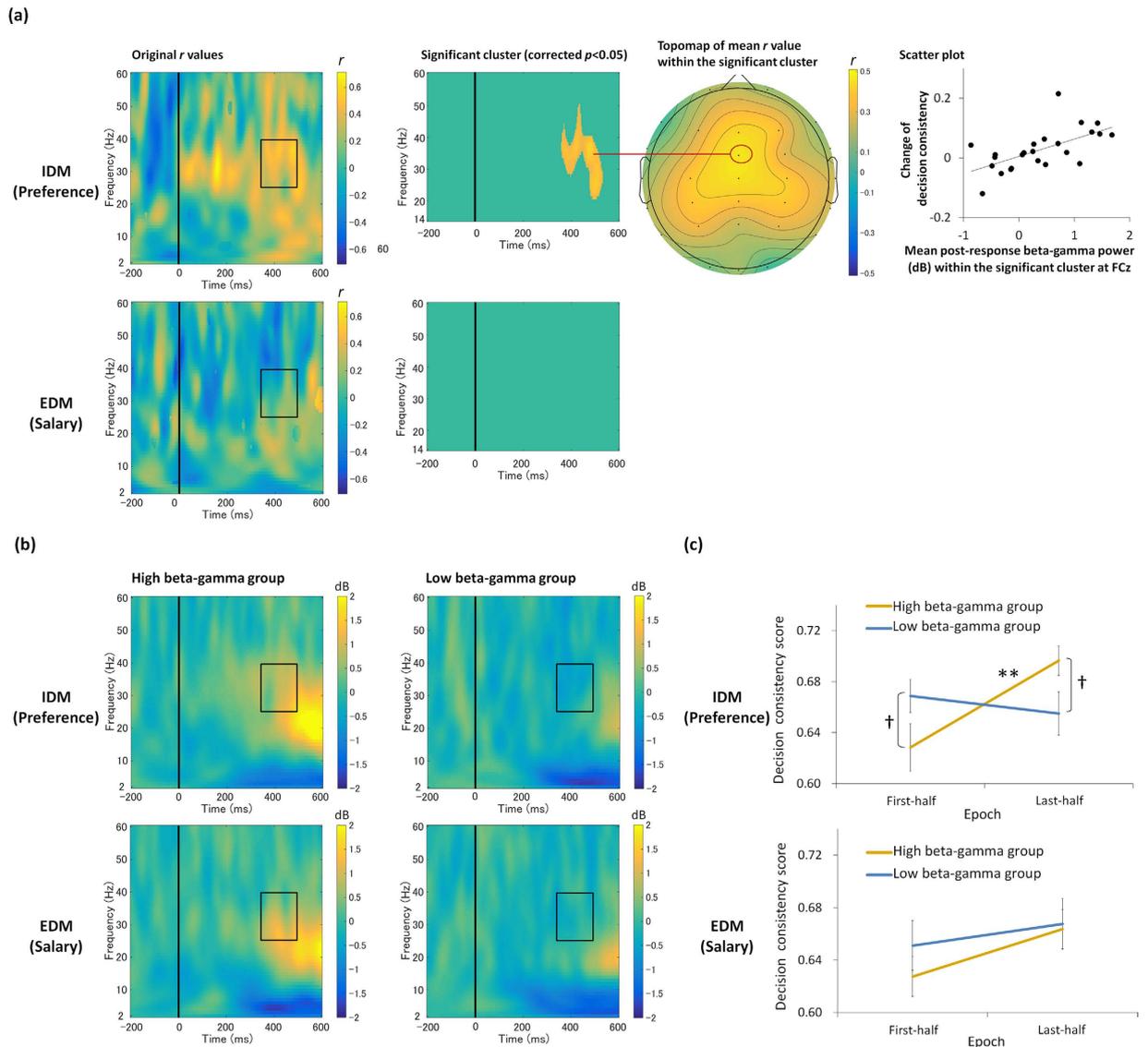


Figure 2. (a) Correlation results between the response-locked event-related spectral perturbations (ERSP) at FCz for the first-half trials and the change of decision consistency (last-half trials–first-half trials) for each decision-making task (see also Fig. S4 for similar results obtained using pre-stimulus baseline corrected response-locked ERSP data). Scalp topography of the mean r -value within the significant cluster for the preference task, and the scatter plot between the change of decision consistency and mean beta–gamma power within the significant cluster are shown on the right side. (b) ERSP images at FCz for the first-half trials of IDM (preference) and EDM (salary) tasks are shown separately for high and low beta–gamma groups of each task. These figures are presented for illustration purposes. (c) Decision consistency for the first-half and the last-half trials in IDM (preference) and EDM (salary judgment) tasks shown separately for high and low beta–gamma groups. †Denotes a marginal difference of $p < 0.10$. **Denotes a significant difference of $p < 0.005$. Error bars represent standard errors. Rectangles presented in the original r -values for preference task (a) and ERSP images (b) show the time–frequency window used for calculating the average power for dividing high and low beta–gamma groups. Internally guided and externally guided decision-making are denoted respectively as IDM and EDM.

Relation between the change of decision consistency and post-response beta gamma power.

Figure 2(a) presents correlation results between the event-related spectral perturbations (ERSP) in the first-half trials at FCz and the change of decision consistency in each task. Positive correlation was found in the beta–gamma (25–40 Hz) range after response in IDM (preference judgment), but not in EDM (salary judgment).

Permutation tests, which are used to avoid multiple comparisons in the large time–frequency space (see *Methods* and Supplemental Information for more details), yield significant positive correlation between the beta–gamma power around 400 ms after response in the IDM task (cluster r -value sum = 271.67, cluster count = 582, corrected $p < 0.05$). This result indicates that participants who showed greater beta–gamma power around 400 ms relative to the pre-response baseline during the first-half trials show more increased decision consistency in last-half trials than in first-half trials. No such significant correlation was found for the EDM task.

We conducted the following additional analysis to ascertain whether the increase of decision consistency in the IDM task is observed in participants who showed increased post-response beta–gamma power in the first-half trials, or not. We divided participants into two groups (12 participants for each group) based on the median of the averaged post-response beta–gamma power (350–500 ms and 25–40 Hz) at FCz (see Fig. 2(b)). Figure 2(c) presents the respective decision consistency scores of the IDM task for these two groups in the first-half and the last-half trials. Two-way mixed ANOVA (groups of beta–gamma power in the first-half trials (high beta–gamma group, low beta–gamma group) \times two epochs (first-half and last-half trials)) revealed a significant main effect of epoch ($F(1, 22) = 5.604, p = 0.03, \eta_G^2 = 0.07$), and significant interaction ($F(1, 22) = 12.75, p = 0.002, \eta_G^2 = 0.14$). Shaffer post-hoc test results revealed that the decision consistency was higher in last-half trials than in the first-half trials in the high beta–gamma group ($p = 0.004$). No such difference was found in the low beta–gamma group. In the first-half trials, the decision consistency was marginally lower in the high beta–gamma group than the low beta–gamma group ($p = 0.09$). In the last-half trials, the decision consistency was marginally higher in the high beta–gamma group than the low beta–gamma group ($p = 0.06$). These results indicate that the increase of decision consistency in the IDM task was observed only in the participants who showed increased post-response beta–gamma power in the first-half trials. Although we conducted the same comparisons for the EDM task by dividing the high and the low beta–gamma groups (12 participants for each group), no significant difference was found (see Fig. 2(b,c)).

Regarding the results of averaged ERSP across all participants, see Supplemental Information (Supplementary Fig. S5).

The gamma power is known to show cross-frequency coupling with theta–alpha phase^{55,56}. Therefore, we examined cross-frequency phase-amplitude coupling between the beta–gamma power and lower theta–alpha phase at FCz. Results obtained by application of permutation-corrected $p < 0.05$ show no significant coupling (see Supplemental Information and Fig. S6 for methods and detailed results).

Discussion

This study was conducted to compare IDM and EDM within the same paradigm in terms of CBL and underlying neural processes. Regarding CBL measured as the increase of decision consistency, significantly higher decision consistencies in the last-half trials than in the first-half trials were found for both IDM (i.e., occupation preference judgment) and EDM (i.e., salary judgment) tasks (Fig. 1(d)), as we had expected. These results were consistent with those obtained in previous studies of CBL conducted using the perceptual-decision-making task⁴⁶ and the free choice paradigm^{39,41,45}.

Differences of CBL between IDM and EDM were observed in relation to the underlying neural activities (Fig. 2). In IDM, participants with increased fronto-central beta–gamma power after response (approximately 400 ms) in the first-half trials showed increased decision consistency. The post-response beta–gamma power predicted the subsequent change of decision consistency as the measurement of CBL. No such relation was found for EDM. These data are the first demonstrating a difference of CBL between IDM and EDM in terms of the beta–gamma activities after response. The post-response beta–gamma is thought to reflect a key process in CBL of IDM.

Then, what is the difference between the underlying mechanisms of IDM and EDM? One plausible difference is whether a reward-related process is recruited or not. Regarding IDM, correlation between the beta–gamma power and the change of decision consistency might reflect the reward-based learning process. As described above, the front–central beta–gamma power after positive feedback is thought to reflect positive reinforcement processes^{16–24}. Although the beta–gamma power in the present data was observed after response instead of positive feedback, the time range and scalp distribution of the present beta–gamma activity show similar physiological properties to those of the reward-based decision making tasks^{16–24}. This evidence implies that reward-based learning process takes part in the CBL of IDM. This interpretation seems reasonable given that the neuroimaging studies of CBL have reported activation within neural substrates related to reward-based reinforcement learning^{40,49–52}. In contrast, regarding EDM, Akaishi *et al.*⁴⁶ reported that CBL is probably mediated by a non-reward based learning process. Considering that the beta–gamma power is related to the positive reinforcement process^{16–24}, the present results of EDM (i.e., no significant relation between the beta–gamma power and CBL) are consistent with results reported by Akaishi *et al.*⁴⁶, who reported no strong activation within the reward-related neural substrates.

When we calculated the index of the CBL at the trial level (trial-based decision consistency; see Supplemental Information including Figs S2 and S3(a)), a change of decision consistency was observed in the IDM but not in EDM. This behavioral result also suggests that the underlying mechanisms of CBL differ between IDM and EDM. Different from the results of epoch-based decision consistency (Last-half–first half; Fig. 1(d)), the increased decision consistency was observed only in the IDM. A possible reason for the discrepancy is that trial-based decision consistency can be affected strongly by the combination of stimulus: if participants selected “Lawyer” in the trial of “Lawyer” vs. “Designer” for preference judgment, the choice would be affected if the more preferred or less preferred occupation were paired with “Lawyer”, whether the “Lawyer” is chosen or rejected at the next opportunity. The combination of the two options was determined randomly across trials. Therefore, if no CBL exists, then no change of decision consistency is observed (this result was confirmed by the Simulation 1 reported in Supplemental Information and Fig. S1(a,b)). However, that effect from the combination of the two options functions as noise to decrease the statistical power to observe decision consistency (both in the epoch-based and trial-based decision consistency). Moreover, trial-based decision consistency would be affected strongly by that noise because fewer trials were used to calculate one index value. It is noteworthy that, even with such trial-based decision consistency, increased decision consistency remained for IDM. This result reflects that the CBL in IDM, which might be induced by reward-based learning, is robust compared to that in EDM, which is thought to be induced by non-reward based learning.

The increase of decision consistency in the IDM task was observed only for participants who showed increased post-response beta–gamma power in the first-half trials (Fig. 2(c)). This result suggests that performing IDM is insufficient for inducing CBL, and suggests that the increased beta–gamma power after response is necessary. Not only external factors (e.g., type of decision-making task, IDM or EDM) but also the internal factors of participants (e.g., individual difference in reactivity of reward-related process for the specific IDM task, which is reflected in the beta–gamma power) can affect the degree of CBL in IDM. Izuma and Murayama⁴² reported from their meta-analysis that the effect size of choice-induced preference change was small. Our results imply that a possible reason for the small effect size is the individual difference of the reactivity for the specific preference task and stimulus. Further studies must be conducted to examine internal factors reflected in the beta–gamma activity to elucidate CBL in IDM.

Despite the importance of our data for revealing the relations between the beta–gamma powers and CBL in IDM, these findings leave several questions unresolved. First, we specifically examined beta–gamma activities based on the notion that the beta–gamma reflects positive reinforcement processes in reward-based learning tasks^{16–22}. Moreover, the reward-related neural substrates are known to be activated in relation to the choice-induced preference change^{40,49–52}. However, because we did not manipulate and measure the amount of reward during the IDM (preference) task, we cannot conclude that the beta–gamma power observed in the present study reflects the reward-related activities as it did in the case of the reward-based reinforcement learning task. Additional studies must be conducted to clarify the functional role of beta–gamma power at the neuronal and psychological levels.

Second, we used the change of decision consistency as the measure of CBL because that index can avoid the methodological flaw of the free choice paradigm^{42,44,54} and because it is applicable for both IDM and EDM tasks. However, we cannot judge whether the increase of decision consistency results from increasing the value of chosen items or decreasing the value of rejected items. Choice-induced preference change is well known to be observed as an increased preference for chosen items and/or decreased preference for rejected items^{39–42}. The increase of decision consistency is visible by one or both possibilities (see Fig. S1(c–e)). To specify the relation between the reinforcement process and beta–gamma power in the CBL of IDM further, it is necessary to conduct studies to measure the index of choice-induced preference change (i.e., change of subjective preference rating for each item) without suffering from the methodological flaw of the free-choice paradigm^{42,44,54}.

Third, we used salary judgment as the EDM task by confirming that the participants' decisions were consistent with the external criteria (i.e., actual average annual salary based on a statistical survey by the Ministry of Health, Labour and Welfare of Japan; see Fig. 1(b)). Nevertheless, the salary task differed from the EDM tasks used in earlier studies (i.e., reward based^{8–10,13} and perceptual decision making⁴⁶). Therefore, one must be careful when generalizing the present result to any kind of EDM task.

Fourth, using simulations (Fig. S1(c–e)) we confirmed an increase of decision consistency when CBL occurs. However, we cannot exclude the possibility that other models can explain the increase of decision consistency better than the model used for the present simulation (Simulation 2, see Supplemental Information). For example, a lower decision noise level in the last-half trials than in the first-half trials can increase the decision consistency, even if no true preference change occurs. This possibility is apparent in our simulation presented in Fig. S1(a): higher decision consistency was observed where less decision noise was included. Although the decision noise level was constant in that simulation, the noise level can be decreased in the last-half trial. Reduction of the decision noise can be interpreted as reflecting that people become able to recognize their own preferences well via a series of decisions. This can be regarded as a kind of CBL. However, this point differs from the explanation assumed for previous CBL studies^{39–41,46,51,52}. It remains unclear which of the true preference changes or reduction of decision noise is the better model for the CBL.

In summary, we first compared CBL in IDM and EDM within the same paradigm. Differences of CBL between these two types of decision making were observed in relation to the underlying neural activities. The beta–gamma power around 400 ms after response predicted the following change of decision consistency in IDM, not in EDM. Although the timing of reinforcement process in EDM (i.e., reward-based decision making) has been well identified, it has been less than clear in IDM. This study provides the first evidence that neural activities after free choice response are related to CBL in IDM. The post-response front-central beta–gamma power is thought to reflect a key process of learning based on one's own decision in IDM.

Methods

Participants. This study examined 24 healthy undergraduate students (11 male; age range = 18–21 years, mean age = 19.58 years) recruited from Hiroshima University. All participants were native Japanese speakers, right-handed, with normal or corrected-to-normal vision. All were free of neurological and psychiatric disorders. No participant was either medicated or a habitual drinker or smoker. All experimental protocols were conducted in accordance with guidelines approved by the Ethical Committee of the Graduate School of Education, Hiroshima University. Written informed consent was obtained from each participant before the investigation. Each participant was paid a small fee for participation.

Stimuli and Tasks. As in previous studies^{12,31}, 28 occupation-related terms (e.g., lawyer, carpenter) were collected from the Classification of Occupation for Employment Security Service (ESCO: <http://www.jil.go.jp/institute/seika/shokugyo/sakuin/>). One term was chosen from each of the third categories of the ESCO. To determine the correct response of the EDM (salary judgment) task, the average annual salary for each occupation was collected from Nensyu-labo (http://nensyu-labo.com/2nd_syokugyou.htm), the salary data of which were based on a statistical survey by the Ministry of Health, Labour and Welfare of Japan. The average annual salary did not covariate with word-length and used-frequency of each occupation-related term (for details, see Supplemental Information). The term pairs for the IDM (occupation preference judgment) and the EDM (salary judgment)

tasks were produced. For each participant and task, 112 pairs were generated randomly with the restriction that each term was used eight times.

In the IDM (occupation preference judgment) task, the two occupation words were presented. Then participants were asked to judge which occupation they would rather do (“Which occupation would you rather do?”) by pressing the button on the corresponding side, as in earlier studies^{12,34} (see Fig. 1(a)). Participants were instructed explicitly that no objectively correct answer exists: they must make their own decisions.

In the EDM (salary judgment) task, participants were asked to judge which is a highly paid occupation on average (“Which occupation is highly paid?”) by pressing the button on the corresponding side (see Fig. 1(a)). Participants were instructed clearly that the average salary is based on the statistical survey by Ministry of Health, Labour and Welfare, and that there was one objective correct answer. Following the method described by Akaishi *et al.*⁴⁶, we did not present the visual feedback in the EDM task to examine the CBL.

Details of the procedures used for these two tasks are presented in the following sections describing procedures.

Procedure. When participants arrived in the laboratory, the experimental procedure was explained. They read an information sheet and signed the consent form. After assigning ratings for each occupation term (for details, see Supplemental Information) and electrode placement, participants were seated on a comfortable chair facing a computer screen in a quiet electrically shielded room. A chin rest was used to help participants maintain the head position during recording.

Participants performed counterbalanced tasks of two types. Four blocks of 28 trials were conducted for each task (see Fig. 1(a)). The presentation side of words was randomized across participants. The order of trials was also randomized. Before the experimental trials, participants were given three practice trials for each task to familiarize them with the tasks.

Each block began with the appearance of an instruction related to the task type on the screen for 3,000 ms. After a 1,000 ms blank, trials began: the fixation cross was presented for 1,000 ms. Then two stimulus words and a question (“Which occupation would you rather do?” or “Which occupation is highly paid?”) were presented. The stimuli and question remained visible on the screen until 1,500 ms after the participant pressed the button. The stimuli and question were replaced with the blank screen, which was presented for 1,500 ms. While this blank screen was displayed, participants were allowed an eye blink. After this blank-screen period, the subsequent trial began (the fixation cross was presented for 1,000 ms). The reaction time (RT) from the presentation of the stimuli to the response was recorded. Participants were instructed to press either the left or right button with the corresponding index finger as quickly and accurately as possible after each stimulus was presented. Additionally, they were asked to avoid eye blinking during times other than the blank screen.

Decision consistency score. To measure the degree of CBL, the decision consistency scores were calculated, respectively, for the first-half and the last-half trials (see Fig. 1(c)). This index can avoid the methodological flaw of the free choice paradigm (see also Fig. S1(a,b), Simulation 1)^{42,44,54}. In addition, the index is applicable for both the IDM and EDM tasks. The decision consistency score represents the rate of trials in which a certain occupation word was repeatedly chosen or rejected. In cases where a participant chose an occupation in a certain trial and it was chosen again in the trial in which the occupation was presented the next time, we counted that trial as a consistent decision. In addition, in cases where a participant rejected an occupation in a certain trial and it was rejected again in the trial in which the occupation was presented the next time, we counted that trial as a consistent decision. The consistent decisions were counted for each occupation word. Then that number was converted to a rate of consistent decision by dividing the number of consistent decisions by the sum of the number of consistent and inconsistent decisions. The average of the rate of consistent decisions was calculated across all occupation words. That figure was used as the decision consistency score.

EEG recordings and analyses. EEG was recorded using 30 silver–silver chloride cup electrodes attached to an electrocap (Quik-Cap; NeuroScan). The signals were amplified with a bandpass of 0.0159–120 Hz and were digitized at a 1,000 Hz sampling rate using the EEG recorder (EEG-1100; Nihon Kohden Corp., Tokyo, Japan). *Supplemental Information* provides more details of EEG recording.

EEG data analysis was conducted using EEGLAB toolbox⁵⁷ running under Matlab 8.4.0 (The Mathworks Inc.). Data were filtered using a low-pass filter of 120 Hz and a high-pass filter of 1 Hz. Response-locked data epochs starting from 1,000 ms before and 1,500 ms after the response button press were extracted. Baselines for event-related potential (ERP) were taken from –500 ms to –250 ms relative to the response onset. Independent component analysis (ICA) was used for artifact rejection from EEG data (for details, see Supplemental Information). After these artifact rejection processes, each type of data epoch was divided into experimental conditions.

Event-related spectral perturbation (ERSP) is the degree to which the spectral power differs from the mean baseline power as a function of time and frequency. For ERSP calculation, the Morlet wavelet was used. Starting from –442 ms before and ending up to 943 ms following response onset, 59 linear-spaced frequencies ranging from 2 Hz to 60 Hz were calculated every 3 ms. The wavelet cycle was increasing from 2 cycles at 2 Hz to 8 cycles at 60 Hz. The pre-response duration (–442–0 ms) was used as the baseline as reported by Wang *et al.*⁵⁸ because CBL is expected to be related to the change of neural activity from before to after decision making. We avoided using the pre-stimulus baseline for results reported in the main text because post-response activities (i.e., the target activities of the present study) of the preceding trial can contaminate the pre-stimulus baseline⁵⁹. However, to confirm the correlation result observed in the post-response duration was not caused by the pre-response baseline activity, this report presents the correlation results obtained after applying the pre-stimulus baseline to the response-locked ERSP in Supplemental Information (Fig. S4).

Regarding statistical analyses of group averaged ERSP, see Supplemental Information.

Correlation analyses. To ascertain whether the participants who showed greater beta–gamma power during the first-half trials show a stronger CBL effect or not (i.e., increased decision consistency in last-half trials than first-half trials), the difference of decision consistency (last half minus first half) was calculated for each participant and task. The response-locked ERSPs of first-half trials of each task at FCz were used for these analyses. Data falling outside mean ± 3 SD were regarded as outliers, and were removed. We performed cluster-based permutation tests⁶⁰ to avoid issues related to multiple comparisons in the large time–frequency space. For the permutation test, the frequency range and time window were limited, respectively, to the beta–gamma band (14–60 Hz) and –200 to 600 ms (for details see Supplemental Information).

References

1. Doya, K. Modulators of decision making. *Nat. Neurosci.* **11**, 410–416 (2008).
2. Dixon, M. L. & Christoff, K. The Lateral Prefrontal Cortex and Complex Value-Based Learning and Decision Making. *Neurosci. Biobehav. Rev.* 1–10, doi: 10.1016/j.neubiorev.2014.04.011 (2014).
3. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
4. Coricelli, G., Dolan, R. J. & Sirigu, A. Brain, emotion and decision making: the paradigmatic example of regret. *Trends Cogn. Sci.* **11**, 258–265 (2007).
5. O’Doherty, J. P., Hampton, A. & Kim, H. Model-based fMRI and its application to reward learning and decision making. *Ann. N. Y. Acad. Sci.* **1104**, 35–53 (2007).
6. O’Doherty, J. P. The problem with value. *Neurosci. Biobehav. Rev.* **43**, 259–268 (2014).
7. Bai, Y., Katahira, K. & Ohira, H. Dual learning processes underlying human decision-making in reversal learning tasks: functional significance and evidence from the model fit to human behavior. *Front. Psychol.* **5**, 1–8 (2014).
8. Katahira, K., Fujimura, T., Matsuda, Y.-T., Okanoya, K. & Okada, M. Individual differences in heart rate variability are associated with the avoidance of negative emotional events. *Biol. Psychol.* **103**, 322–331 (2014).
9. Bai, Y., Katahira, K. & Ohira, H. Valence-separated representation of reward prediction error in feedback-related negativity and positivity. *Neuroreport* **26**, 157–162 (2015).
10. Glascher, J., Hampton, A. N. & O’Doherty, J. P. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb. Cortex* **19**, 483–495 (2009).
11. Nakao, T., Ohira, H. & Northoff, G. Distinction between externally vs. internally guided decision-making: Operational differences, meta-analytical comparisons and their theoretical implications. *Front. Neurosci.* **6**, 1–26 (2012).
12. Nakao, T., Bai, Y., Nashiwa, H. & Northoff, G. Resting-state EEG power predicts conflict-related brain activity in internally guided but not in externally guided decision-making. *Neuroimage* **66**, 9–21 (2013).
13. O’Doherty, J., Critchley, H., Deichmann, R. & Dolan, R. J. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J. Neurosci.* **23**, 7931–7939 (2003).
14. Tanaka, S. C. *et al.* Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* **7**, 887–893 (2004).
15. Kim, H., Shimojo, S. & O’Doherty, J. P. Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol.* **4**, e233 (2006).
16. Doñamayor, N., Schoenfeld, M. A. & Munte, T. F. Magneto- and electroencephalographic manifestations of reward anticipation and delivery. *Neuroimage* **62**, 17–29 (2012).
17. Doñamayor, N., Marco-Pallarés, J., Heldmann, M., Schoenfeld, M. A. & Münte, T. F. Temporal dynamics of reward processing revealed by magnetoencephalography. *Hum. Brain Mapp.* **32**, 2228–2240 (2011).
18. Hosseini, A. H., Rodríguez-Fornells, A. & Marco-Pallarés, J. The role of beta-gamma oscillations in unexpected rewards processing. *Neuroimage* **60**, 1678–1685 (2012).
19. Marco-Pallarés, J., Münte, T. F. & Rodríguez-Fornells, A. The role of high-frequency oscillatory activity in reward processing and learning. *Neurosci. Biobehav. Rev.* **49**, 1–7 (2014).
20. Mas-Herrero, E., Ripollés, P., HajiHosseini, A., Rodríguez-Fornells, A. & Marco-Pallarés, J. Beta Oscillations and Reward Processing: Coupling Oscillatory Activity and Hemodynamic Responses. *Neuroimage* **119**, 13–19 (2015).
21. Marco-Pallares, J. *et al.* Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia* **46**, 241–248 (2008).
22. Jung, J. *et al.* Brain responses to success and failure: Direct recordings from human cerebral cortex. *Hum. Brain Mapp.* **31**, 1217–1232 (2010).
23. Cohen, M. X., Elger, C. E. & Ranganath, C. Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage* **35**, 968–978 (2007).
24. Feingold, J. *Beta oscillations in frontal cortex and striatum represent post-processing of successful behavior.* (Harvard University: MIT Division of Health Sciences and Technology, 2011).
25. Goldberg, E. & Podell, K. Adaptive versus Veridical Decision Making and the Frontal Lobes. *Conscious. Cogn.* **8**, 364–377 (1999).
26. Goldberg, E. & Podell, K. Adaptive Decision Making, Ecological Validity, and the Frontal Lobes. *J. Clin. Exp. Neuropsychol.* **22**, 56–68 (2000).
27. Lieberman, M. D. & Eisenberger, N. I. In *On building, defending and regulating the self: A psychological perspective* (eds. Tesser, A., Wood, J. V. & Stapel, D. A.) 77–102 (Psychology Press, 2005).
28. Mograbi, G. J. Neural basis of decision-making and assessment: issues on testability and philosophical relevance. *Mens Sana Monogr.* **9**, 251–259 (2011).
29. Nakao, T. *et al.* The degree of early life stress predicts decreased medial prefrontal activations and the shift from internally to externally guided decision making: An exploratory NIRS study during resting state and self-oriented task. *Front. Hum. Neurosci.* **7** (2013).
30. Volz, K. G., Schubotz, R. I. & von Cramon, D. Y. Decision-making and the frontal lobes. *Curr. Opin. Neurol.* **19**, 401–406 (2006).
31. Nakao, T. *et al.* Self-Knowledge Reduces Conflict by Biasing One of Plural Possible Answers. *Personal. Soc. Psychol. Bull.* **36**, 455–469 (2010).
32. Johnson, S. C. *et al.* The cerebral response during subjective choice with and without self-reference. *J. Cogn. Neurosci.* **17**, 1897–1906 (2005).
33. Knutson, B. *et al.* Neural antecedents of the endowment effect. *Neuron* **58**, 814–822 (2008).
34. Di Domenico, S. I., Rodrigo, A. H., Ayaz, H., Fournier, M. A. & Ruocco, A. C. Decision-making conflict and the neural efficiency hypothesis of intelligence: A functional near-infrared spectroscopy investigation. *Neuroimage* **109**, 307–317 (2015).
35. Di Domenico, S. I., Fournier, M. A., Ayaz, H. & Ruocco, A. C. In Search of Integrative Processes: Basic Psychological Need Satisfaction Predicts Medial Prefrontal Activation During Decisional Conflict. *J. Exp. Psychol. Gen.* **142**, 967–978 (2013).
36. Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M. & Cohen, J. D. The neural bases of cognitive conflict and control in moral judgment. *Neuron* **44**, 389–400 (2004).

37. Kahane, G. *et al.* The neural basis of intuitive and counterintuitive moral judgment. *Soc. Cogn. Affect. Neurosci.* **7**, 393–402 (2012).
38. Moll, J. *et al.* Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. USA* **103**, 15623–15628 (2006).
39. Izuma, K. *et al.* A causal role for posterior medial frontal cortex in choice-induced preference change. *J. Neurosci.* **35**, 3598–3606 (2015).
40. Izuma, K. *et al.* Neural correlates of cognitive dissonance and choice-induced preference change. *Proc. Natl. Acad. Sci. USA* **107**, 22014–22019 (2010).
41. Nakamura, K. & Kawabata, H. I Choose. Therefore I Like: Preference for Faces Induced by Arbitrary Choice. *PLoS One* **8**, e72071 (2013).
42. Izuma, K. & Murayama, K. Choice-induced preference change in the free-choice paradigm: a critical methodological review. *Front. Psychol.* **4**, 41 (2013).
43. Brehm, W. J. Postdecision changes in the desirability of alternatives. *J. Abnorm. Psychol.* **52**, 384–389 (1956).
44. Alós-ferrer, C. Choice-induced preference change and the free-choice paradigm: A clarification. *Judgm. Decis. Mak.* **10**, 34–49 (2015).
45. Koster, R., Duzel, E. & Dolan, R. J. Action and Valence Modulate Choice and Choice-Induced Preference Change. *PLoS One* **10**, e0119682 (2015).
46. Akaishi, R., Umeda, K., Nagase, A. & Sakai, K. Autonomous Mechanism of Internal Choice Estimate Underlies Decision Inertia. *Neuron* **81**, 195–206 (2014).
47. Niv, Y., Langdon, A. & Radulescu, A. A free-choice premium in the basal ganglia. *Trends Cogn. Sci.* **1–2** doi: 10.1016/j.tics.2014.09.005 (2014).
48. Cockburn, J., Collins, A. G. E. & Frank, M. J. A Reinforcement Learning Mechanism Responsible for the Valuation of Free Choice. *Neuron* **83**, 1–7 (2014).
49. Sharot, T., De Martino, B. & Dolan, R. J. How Choice Reveals and Shapes Expected Hedonic Outcome. *J. Neurosci.* **29**, 3760–3765 (2009).
50. Jarcho, J. M., Berkman, E. T. & Lieberman, M. D. The neural basis of rationalization: Cognitive dissonance reduction during decision-making. *Soc. Cogn. Affect. Neurosci.* **6**, 460–467 (2011).
51. Qin, J. *et al.* How choice modifies preference: Neural correlates of choice justification. *Neuroimage* **55**, 240–246 (2011).
52. Kitayama, S., Chua, H. F., Tompson, S. & Han, S. Neural mechanisms of dissonance: An fMRI investigation of choice justification. *Neuroimage* **69**, 206–212 (2013).
53. Valentin, V. V., Dickinson, A. & O’Doherty, J. P. Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* **27**, 4019–4026 (2007).
54. Chen, M. K. & Risen, J. L. How choice affects and reflects preferences: revisiting the free-choice paradigm. *J. Pers. Soc. Psychol.* **99**, 573–594 (2010).
55. Cohen, M. X. *et al.* Good Vibrations: Cross-frequency Coupling in the Human Nucleus Accumbens during Reward Processing. *J. Cogn. Neurosci.* **21**, 875–889 (2008).
56. Fujisawa, S. & Buzsáki, G. A 4 Hz Oscillation Adaptively Synchronizes Prefrontal, VTA, and Hippocampal Activities. *Neuron* **72**, 153–165 (2011).
57. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
58. Wang, L. *et al.* Disentangling the impacts of outcome valence and outcome frequency on the post-error slowing. *Sci. Rep.* **5**, 8708 (2015).
59. Northoff, G., Qin, P. & Nakao, T. Rest–stimulus interaction in the brain: a review. *Trends Neurosci.* **33**, 277–284 (2010).
60. Cohen, M. X. *Analyzing Neural Time Series Data.* (MIT Press, 2014).

Acknowledgements

This research was supported by the Center of Innovation Program of the Japan Science and Technology Agency (JST) and by JSPS KAKENHI Grants 26285168 and 25870467. We thank members of the Faculty of Education, Hiroshima University for data collection.

Author Contributions

T.N. designed the experiment, analyzed the data, and wrote the manuscript. M.O., Y.I., Y.H., R.N., H.O. and R.H. collected the data. N.K., K.K., M.M. and G.N. reviewed and revised the manuscript. All authors approved the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Nakao, T. *et al.* Post-response $\beta\gamma$ power predicts the degree of choice-based learning in internally guided decision-making. *Sci. Rep.* **6**, 32477; doi: 10.1038/srep32477 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016