Electrophysiological mechanisms of foraging decisions and feedback processing

Thesis for the degree of

doctor rerum naturalium (Dr. rer. nat.)

approved by the Faculty of Natural Sciences of Otto von Guericke University Magdeburg

by M.Sc. Franziska Kirsch
born on November 3, 1992 in Schmalkalden

Examiner:
Prof. Dr. med. habil. Markus Ullsperger
Prof. Dr. Christian Bellebaum

submitted on August 29, 2023
defended on March 7, 2024

Zusammenfassung

Die Jobsuche stellt für Menschen einen schwierigen Entscheidungsprozess dar, bei der aktuelle Optionen verglichen, potentielle zukünftige Angebote berücksichtigt und aus Erfahrungen gelernt werden muss um letztlich die richtige Wahl zu treffen. Das Verständnis solch komplexer Entscheidungsfindungsprozesse und Feedback-Mechanismen, die den Erfahrungen im wirklichen Leben ähneln, ist jedoch noch unvollständig. Die sequenzielle Entscheidungsfindung beim Menschen wurde bisher nur begrenzt untersucht. Außerdem gibt es widersprüchliche Erkenntnisse über die zugrunde liegende neuronale Dynamik im Zusammenhang mit Feedbackverarbeitung.

Diese Arbeit umfasst zwei Untersuchungen zur Entscheidungsfindung und Feedback-Verarbeitung. Die erste Studie befasst sich mit dem Explorations-/Exploitationsdilemma und untersucht das Gleichgewicht zwischen der Nutzung vorhandener und der Suche nach neuen Ressourcen. Sie konzentriert sich insbesondere auf Entscheidungen zum sogenannten *Foraging*, bei denen Individuen eine aktuelle Option ausschöpfen und dabei den richtigen Zeitpunkt für die Erkundung neuer Ressourcen bestimmen müssen. Dabei war die Hypothese, dass Individuen sich normativen Vorhersagen über optimales Verhalten in dieser Foraging-Umgebung annähern. Weiterhin wurde erwartet, dass sich das adaptive Entscheidungsverhalten in neuronalen Korrelaten widerspiegelt, die den zeitlichen Verlauf dieser Entscheidung beschreiben. Die zweite Studie befasst sich mit typischen Performanz-rückmeldungsbezogenen elektrophysiologischen Korrelaten (Feedback-related negativity: FRN; P3) und ihrer Kodierungsspezifität in Bezug auf Feedback-Valenz und Erwartbarkeit. Hier lautete die Hypothese, dass sich Belohnungsvorhersagefehler (RPE) im neuronalen Signal nach der Rückmeldung widerspiegeln. Außerdem wurde erwartet, dass der FRN-Latenzbereich einen vorzeichenbehafteten RPE kodiert, der Valenz und Unerwartbarkeit einschließt. Das Signal sollte sich dementsprechend zwischen Feedback, welches besser oder schlechter als erwartet ist, unterscheiden. Das Signal im P3-Latenzbereich sollte hauptsächlich durch Überraschung beeinflusst sein. Darüber hinaus wurde angenommen, dass die Verhaltensanpassung nach Feedback mit dem Signal im FRN-Latenzbereich assoziiert ist.

Um diese Hypothesen zu testen, wurden zwei kognitive Experimente mit großen Stichproben gesunder Teilnehmer durchgeführt und 64-Kanal-EEG-Daten aufgezeichnet. Das erste Experiment beinhaltete eine Aufgabe, bei der die Teilnehmer den Belohnungsgewinn innerhalb eines bestimmten Zeitrahmens maximieren sollten, indem sie Entscheidungen darüber trafen, wann sie die aktuellen Optionen zugunsten von potenziell lohnenderen Alternativen verlassen. Im zweiten Experiment nahmen die Teilnehmer an einer Zeitschätzungsaufgabe teil, bei der sie auf der Grundlage von manipulierten Zeitfenstern für richtige Antworten ein erwartetes oder unerwartetes Feedback erhielten.

Anhand von Regressionsanalysen wurde festgestellt, dass die Teilnehmer sich normativen Entscheidungen bei der Nahrungssuche annäherten, aber systematische Verzerrungen aufwiesen. Es wurden unterschiedliche Frequenzmuster in den Delta-, Theta- und Beta-Bändern identifiziert, die mit Explorations- und Exploitationsmechanismen in Foraging-Paradigmen verbunden sind. In Bezug auf rückmeldungsbezogene Prozesse wurden bestehende Befunde im FRN-Latenzbereich repliziert, die einen vorzeichenbehafteten RPE widerspiegeln, während der P3-Latenzbereich mit Überraschung und positivem Feedback in Verbindung gebracht wurde. Bemerkenswerterweise korrelierte der P3-Latenzbereich, und nicht die FRN, mit Verhaltensanpassungen nach dem Feedback.

Insgesamt deuten diese Ergebnisse auf ein komplexes Zusammenspiel verschiedener neuronaler Oszillationen hin, die für die sequenzielle Entscheidungsfindung verantwortlich sind. Der Entscheidungsprozess beginnt früh in der Foraging-Phase und kumuliert möglicherweise, bis ein Schwellenwert erreicht wird, der zu einer Entscheidung führt. In künftigen Studien könnte die Untersuchung der Rolle etablierter rückmeldungsbezogener Korrelate für das Lernen in diesen Entscheidungsprozessen wertvolle Erkenntnisse liefern.

Abstract

Job search represents a difficult decision-making process for people, where current options have to be compared, potential future offers have to be considered, and learning from experience is necessary to ultimately make the right choice. However, the understanding of such complex decision-making processes and feedback mechanisms similar to real-life experiences remains incomplete. Sequential decision-making in humans has received limited attention, and there is conflicting evidence regarding the underlying neuronal dynamics related to feedback.

This work aims to address this issue within two investigations on decision-making and feedback processing. The first study addressed the exploration-exploitation dilemma. Specifically, it focused on foraging decisions, where individuals must exploit a current option while determining the right time to explore new resources. I expected that individuals approximate normative predictions of optimal behavior within this foraging environment, reflected in neural correlates that describe the temporal progression of this decision. The second study delved into typical feedback-related electrophysiological correlates (Feedback-related negativity: FRN; P3) and their encoding specificity concerning feedback valence and expectedness. The hypothesis suggested that reward prediction errors (RPE) are coded in the neural signal post-feedback. The FRN latency range was thought to encode a signed RPE, incorporating valence and unexpectedness, whereas the P3 should mainly be influenced by surprise. Moreover, behavioral adaptation was expected to be associated with the signal in the FRN latency range.

To test these hypotheses, two cognitive experiments were conducted with large samples of healthy participants, recording 64-channel EEG data. The first experiment involved a patch-leaving task, where participants aimed to maximize reward gain within a certain time frame, making decisions on when to leave current options for potentially more rewarding alternatives. In the second experiment, participants engaged in a time estimation task, receiving feedback based on manipulated correct answer time windows to introduce expected or unexpected feedback.

Results show that participants approximated normative foraging decisions but exhibited systematic biases. Through single-trial regression analyses, I identified distinct frequency patterns in the delta, theta, and beta bands, linked to exploration and exploitation mechanisms in patch-leaving paradigms. Concerning feedback-related processes, existing findings of the EEG signal in the FRN latency range reflecting a signed RPE were replicated, while the P3 latency range was associated with surprise and positive valence feedback. Notably, the P3 latency range, rather than the FRN, correlated with post-feedback behavioral adjustments.

Overall, these results indicate a complex interplay of various neural oscillations responsible for sequential decision-making, starting early in the foraging phase and possibly accumulating until a threshold is reached, leading to a decision. In future studies, exploring the role of established feedback-related correlates on learning in these decision-making processes could yield valuable insights.

List of abbreviations

ACC	anterior cingulate cortex
aMCC	anterior midcingulate cortex
BRR	background reward rate
dACC	dorsal anterior cingulate cortex
EEG	electroencephalography
ERN	error-related negativity
ERP	event-related potential
FRN	feedback-related negativity
FRR	foreground reward rate
MVT	marginal value theorem
PE	prediction error
PERI	post-error reduction of interference
PES	post-error slowing
PIA	post-error improvements in accuracy
pMFC	posterior medial frontal cortex
RewP	reward positivity
RFL	reinforcement learning
RPE	reward prediction error
vmPFC	ventromedial prefrontal cortex

List of figures

Contents

1 Introduction

Decision-making involves the cognitive process of selecting a course of action from various alternatives. It encompasses the evaluation of options based on their potential outcomes and choosing the option that is most likely to achieve a desired result. This fundamental aspect of human behavior is present in numerous daily life situations, ranging from simple choices like what to have for breakfast to complex decisions such as career paths or purchasing a house.

Rangel et al. (2008) proposed a decision-making cycle comprising several stages (***Figure 1-1***). Initially, potential actions are represented in relation to internal and external states. These actions are then evaluated given the specific states, and their values are compared to make a decision on an action. Once a decision is made, the desirability of the outcome is assessed by experiencing its consequences and calculating prediction errors. In this work, various aspects of this decision-making cycle are investigated in the context of two studies, aiming to enhance our understanding of the cognitive processes involved in decision-making among healthy adults.

The first study primarily focuses on examining the stages of valuation and decision-making, specifically within the context of a sequential decision-making task known as foraging. The second study explores outcome evaluation and, to some extent, learning processes. It particularly investigates how individuals respond to feedback and engage in feedback-based learning. In the subsequent chapters, I provide the theoretical background for both the valuation and decision-making study and the evaluation and learning study.

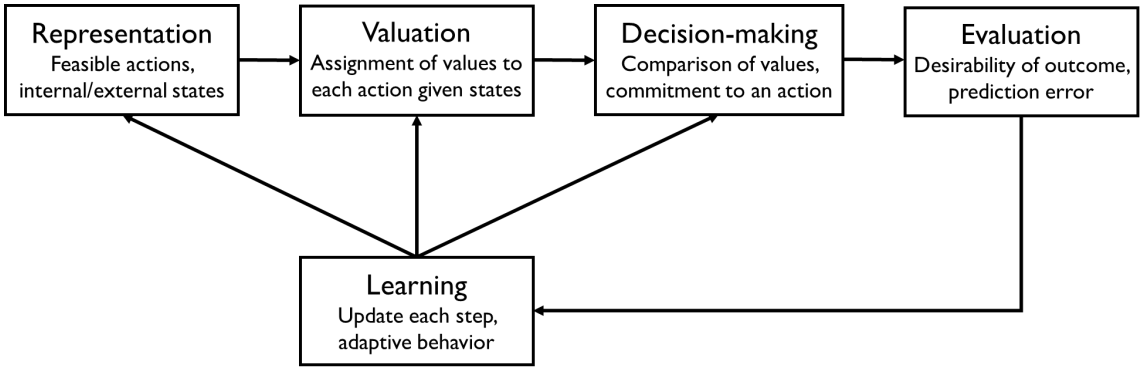

Figure 1-1. The cycle of decision-making. Adaptive behavior comprises five consecutive steps: representation, valuation, decision-making, outcome evaluation, and learning. Initially, the agent forms representations of feasible actions (e.g., continue eating from the current berry bush vs. switching a berry bush) and environmental (e.g., availability of berry bushes) or internal states (e.g. feeling hungry). Valuation involves assigning expected values to actions (e.g., high vs. low expected reward and effort of staying vs. switching to a far bush). Decision making entails comparing the values of actions and selecting the one with the highest value (e.g., staying at the current bush, because value of staying outweighs value of switching in terms of energy). Outcome evaluation assesses received outcomes compared to expected values (e.g., comparison of expected reward gain from staying with actual gain), driving learning (e.g., adjust behavior and switch if the reward obtained from current bush diminishes) and updates to representations, expected values, and decision-making strategies (e.g., exploration-exploitation trade-off). This iterative process enables the agent to adapt its behavior based on experience and feedback, promoting effective interaction with the environment. This scheme is adapted from Rangel et al. (2008).

1.1 Optimal sequential decision making

As mentioned earlier, decision-making is usually described as the ability to learn about the relative value of various available options, draw conclusions from this and choose the best option (Jocham et al., 2011). This gave rise to many studies that investigated decision-making represented by the choice between two or more distinct options and outlined neural correlates for it (Busemeyer et al., 2019; Frömer et al., 2019; Gluth et al., 2014; Lee et al., 2021; Polanía et al., 2014). However, decisions in the daily life are usually much more complex than the choice between two more or less rewarding options (Kolling & O'Reilly, 2018). Individuals often face decisions with far-reaching consequences without knowing possible future developments. This type of decision-making comes with the necessity of planning ahead and prospect, but has been largely neglected by previous research (Hunt et al., 2021). Sequential and temporal extended choice paradigms such as foraging paradigms (sometimes also called patch-leaving

decision paradigms) are novel, biology inspired approaches to better understand more complex decisions. Originally, the term foraging refers to the behavior of animals. Imagine a bird searching for grains in a certain meadow. At the beginning there are plenty of grains, but after some time of searching and eating, there are only few grains left. Now the bird has to expend a lot of energy by running around to find the remaining grains. This is when the bird decides to incur the cost of flying to the next meadow to have easier access to new grains again. Additional aspects such as the presence of a possible predator (e.g., a cat) or competition from other birds must be considered by the animal. It is a challenge to find a balance between the energy gained from resources and the energy expended to acquire them in a changing environment. Optimal decision making in such a scenario will ensure the survival of the animal. Indeed, humans also rely on foraging decisions in various aspects of their daily lives. Consider a scenario where you find yourself in a garden with several berry bushes, aiming to maximize your berry harvest within a one-hour time frame. To accomplish this, you must decide on the ideal duration to spend at each bush before transitioning to the next, which is again full of berries. A more complex example is the process of job applications and deciding when to accept an offer without knowing what future opportunities may arise. Achieving satisfactory results here requires the ability to make optimal decisions in volatile environments.

1.1.1 Classification and distinction of decision-making

The following section will illuminate different types and aspects of decision-making and distinguish them from each other.

Value-based decisions are often represented as choices between simultaneously presented concrete offers. Foraging decision-making goes beyond this approach by considering the environmental value, i.e., the value of the entire surroundings, not just the action value of current options. A notable characteristic of foraging is its sequential nature, where individuals encounter specific opportunities and make decisions to accept or reject them. These decisions are based on a comparison between the presented offer and the individual's prediction of what the current or future environments have to offer.

Barack and Gold (2016) suggested different trade-offs that make up the decision-making process: sensitivity to stability versus change for past information, speed versus accuracy for current information, and exploitation versus exploration for future goals. While all aspects can play a crucial role in foraging decisions, the aim of foraging decisions is to minimize energy, time, and costs while maximizing reward in the long term and can therefore be seen as a form of an exploitation-exploration dilemma (Addicott et al., 2017). In the short term, exploitation of current resources maximizes rewards, however, the information obtained during exploration can later be used in the long term. Referring to the aforementioned example, individuals aim to strike a balance between exploring different bushes for berries and exploiting the current bush. Achieving optimal decision-making typically involves an adaptive strategy that combines both exploration and exploitation. In terms of the foraging approach, an exploitation bias could mean staying too long on current and known options while missing out on potentially superior options. An exploration bias on the other hand is characterized by leaving current options too early looking for better options to come.

In some ways, foraging decisions are comparable to perceptual evidence accumulation. Nevertheless, both processes have substantial differences (Polanía et al., 2014). Perceptual evidence accumulation focuses on the integration of sensory information by continuously sampling and accumulating sensory evidence until a decision threshold is reached (Brosnan et al., 2020; T. Liu & Pleskac, 2011; Pereira et al., 2021; Ploran et al., 2007). Here, attentional and action related processes seem to have dissociable roles (Shenhav et al., 2018). Foraging decision-making refers to a broader concept that involves strategic top-down processing on higher levels and encompasses the decisions individuals make to optimize resource acquisition and dynamically adapt to the changing environment. While accumulating evidence might also be a part of the foraging decision process (Davidson & El Hady, 2019; Kristjánsson et al., 2020), there are more complex processes that use evidence accumulation only under certain conditions, and possibly asymmetrically.

To gain insights into human cognition and decision-making, ecological models of animal behavior provide effective tools for describing behavioral adaptations related to optimizing behavior and ensuring survival. Given that humans, like animals, have evolved

to tackle such problems, it is logical to assume that this approach can assist in uncovering the fundamental elements of human foraging functions, especially in the realm of adaptive decision-making. I will discuss this in more detail in the next section.

1.1.2 Optimal foraging theory

The most famous approach in optimal foraging is the Marginal-value-theorem (MVT, Charnov, 1976). It states that the ideal time to leave a current option is when a particular threshold is reached: when the instant reward intake of the current option falls below the average reward intake of the environment (Kolling & Akam, 2017). Here, the instant reward intake of the current option per time is also called foreground reward rate (FRR), while the average intake of the environment per time is called background reward rate (BRR, Gabay & Apps, 2021; Le Heron et al., 2020). Within this framework, the calculation of action utility in energy units is based on the following equation (*(1-1)*: the energy expended to acquire the reward (*e*) is subtracted from the energy gained from the reward (α), divided by total time *T* spent acquiring the reward (Shadmehr & Ahmed, 2020).

$$U = \frac{\alpha - e}{T} \tag{1-1}$$

This theorem makes specific predictions about the ideal leaving time in order to maximize reward (***Figure 1-2***). Various factors can influence the FRR, and among them is the quality of the patch, referring to the yield of available rewards. When the patch offers a higher yield, it results in an initially higher rate of reward acquisition. Consequently, the FRR will take a longer time to reach the background reward rate. Thus, when a patch yields a high reward, staying longer on this patch is beneficial compared to a patch yielding low reward. Contrarily, when the reward history is high, or in other words, the environment is full of rich options, it is beneficial to leave individual patches earlier to exploit as many options as possible compared to a poor environment, where the reward history is low. In that way, instantaneous and average reward rate (FRR and BRR, respectively) independently impact when to leave (Le Heron et al., 2020).

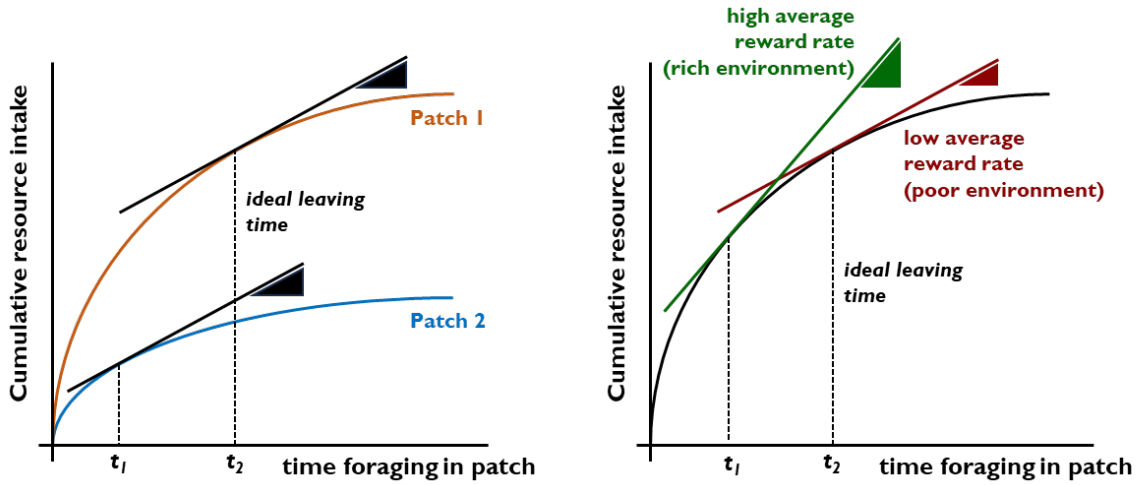

Figure 1-2. Predictions of the Marginal Value Theorem (MVT). In the left figure, it demonstrates that the optimal time to leave a patch with lower rewards is shorter compared to a patch with higher rewards. On the right figure, it reveals that the ideal time to leave the same patch is shorter in a high average reward rate scenario (rich environment) compared to a low average reward rate scenario (poor environment). This scheme is adapted from Shadmehr & Ahmed (2020).

In simple foraging tasks, human behavior aligns closely with the predictions of the MVT. However, ambiguities arise when tasks become more complex, such as situations where reward rates within patches do not consistently decline but can also increase (Kolling & Akam, 2017; Wittmann et al., 2016). In such cases, the optimal strategy may vary, necessitating a stronger weighting of either the past reward rate or the recent reward rate based on the circumstances. Kolling and Akam (2017) argue that to comprehensively capture human foraging behavior in complex environments, a model-free estimation of the overall environmental quality, as suggested by MVT, must be combined with a model-based prediction of forthcoming rewards (see 2.2 for more details). It is still not fully understood, how reward rates are represented in humans, and thus they are often modelled differently. MVT assumes complete information of the environment, therefore the average reward rate is, for example, often defined as the average reward rate of *all* patches in the environment or as the extrapolated future reward rate. On the other hand, because it is impossible in reality to know the environment completely (environmental quality is a probability function and therefore uncertain), the average reward rate is usually approximated by the average rate of the patches that have been foraged so far. Additionally, the MVT uses objective values and assumes that humans always stop exploiting the current patch, as soon as the threshold

is reached. In that way, subjective values and interindividual biases such as risk sensitivity are neglected.

Nevertheless, it has been shown that the MVT model is a good approximation of human behavior and therefore a powerful framework for understanding aspects of patch leaving decisions. The next sections, I will describe in detail the findings to date and what limitations remain.

1.1.3 Foraging decisions in humans

In a patch-leaving paradigm that involved harvesting depleting apple trees through distinct choices, Constantino and Daw (2015) demonstrated that a learning rule proposed by the MVT provided a better explanation of human foraging behavior compared to temporal-difference-error learning. In a study by Wittmann et al. (2016), participants engaged in a task where they repeatedly encountered a patch characterized by stochastic outcomes with an underlying trend of either increasing or decreasing rewards. Results showed that rewards acquired shortly prior to the decision to leave the patch encouraged participants to remain, whereas rewards obtained well in advance of the decision promoted patch leaving. An advanced reinforcement learning model, which estimated reward rate gradients through averaging reward prediction errors, adeptly predicted participants' choices to leave while encompassing these opposing effects of reward trends. The authors showed that a model-free reinforcement learning approach is insufficient for comprehensively explaining adaptable foraging behavior. Instead, it becomes evident that more intricate models must encompass a representation of the environmental value and its dynamic alterations to effectively capture the complexities of such behavior (Kolling & Akam, 2017; Steixner-Kumar & Gläscher, 2020). Le Heron et al. (2020) introduced a continuous patch-leaving paradigm, where participants pressed a button until they decided to leave the current patch. Patches delivered rewards at an exponentially decreasing rate (FRR), varying between patches. The environmental value (BRR) was adjusted by altering the occurrence probabilities of different patch qualities. While manipulations of the FRR und BRR have been shown to independently influence participants' behavior, individuals generally show a tendency to stay longer than an optimal MVT-agent would predict (Hutchinson et al., 2008; Le Heron et al., 2020). Harhen and Bornstein (2023) investigated this bias with the help of a serial

stay-switch task. Their findings suggest that participants learn a representation of the environmental structure through individual patch experiences. The authors argue that uncertainty in environmental structure leads to adaptive discounting of future rewards, resulting in this phenomenon referred to as "overharvesting". Moreover, an asymmetric effect of the BRR was identified: individuals fail to adjust when the environment deteriorates compared to when it improves, leading to suboptimal choices (Garrett & Daw, 2020). In a task where reward magnitude and reward time delay were manipulated, the authors identified the presence of an optimism bias: in deteriorating environments, options that were valuable given the environmental conditions were prematurely discarded in the false anticipation that better options would emerge. While evidence for human foraging behavior and underlying representations is growing, the exact neuro-computational mechanisms that facilitate these foraging choices remain poorly understood. In the next section, I describe neural correlates of decision-making.

1.1.4 Neural correlates and methods of investigating decision-making

Distinct neural mechanisms are present for different types of decision-making processes, such as foraging or the comparison and selection among a limited set of distinct options. The latter has been associated with activity in the ventromedial prefrontal cortex (vmPFC) related to tracking the value of available options (Jocham et al., 2011). In the same study, also striatal prediction error coding predicted the choice performance of participants (***Figure 1-3***).

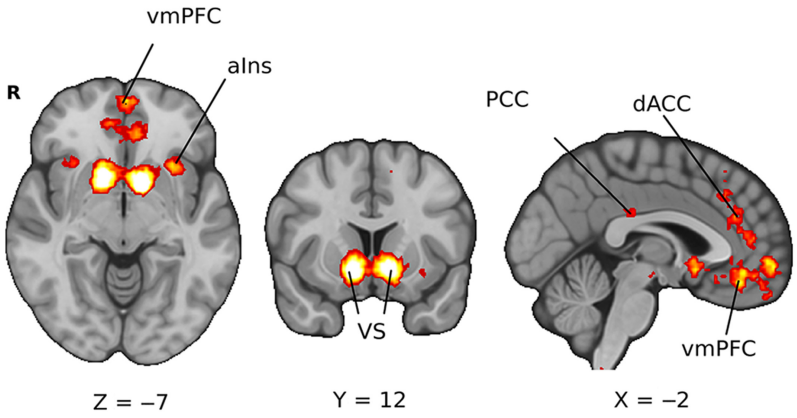

Figure 1-3. Brain regions associated with decision-making. Results of an association test for an automated meta-analysis of value-based decision-making studies provided by the Neurosynth platform are shown. Key regions highlighted include the ventral striatum (VS), ventromedial prefrontal cortex (vmPFC), dorsal anterior cingulate cortex (dACC), posterior cingulate cortex (PCC), and anterior insula (aIns). *Wiley Interdisciplinary Reviews, Dennison, J. B., Sazhin, D., & Smith, D. V., Cognitive Science, Decision neuroscience and neuroeconomics: Recent progress and ongoing challenges, 13(3), e1589, 2021, reproduced with permission from WILEY.*

When it comes to the exploration-exploitation trade-off, there are many studies trying to map neuronal areas to specific responsibilities (for a review, see Mansouri et al., 2017). Opponent processes in frontoparietal regions for undirected and directed exploration, and exploitation are suggested (***Figure 1-4***). While the medial part of the frontopolar region is proposed to be responsible for monitoring the relevance of the current goal and eventually desist from this when the environment is changing, the lateral part is associated with a more elaborated exploration monitoring alternative goals and redirecting cognitive resources to new goals. The posterior prefrontal cortex is supposed to execute the current task, optimizing performance, and thus exploit the current goal. An alternative approach suggests the cooperation of a network of neural regions including the anterior midcingulate cortex (aMCC; sometimes also called dorsal anterior cingulate cortex, dACC) to perform explore-exploit decisions (***Figure 1-3***; Dennison et al., 2022; Donoso et al., 2014). Domenech et al. (2020) was able to study single unit recordings from the prefrontal cortices of epilepsy patients in a volatile environment. Activity in the high-gamma frequency band was found to be responsible for arbitration between exploiting and exploring, while feedback processes associated with exploration showed an increase in the activity in the beta frequency band and a decrease in the theta frequency band and alpha frequency band. Nevertheless, it is important to note that the paradigm employed in this study differed from the typical

foraging paradigm (patch-leaving approach), as it did not require making time-sensitive decisions. Furthermore, the authors of the study focused on investigating feedback-related processes related to the exploration-exploitation dilemma, which diverges from the approach we have taken in our research.

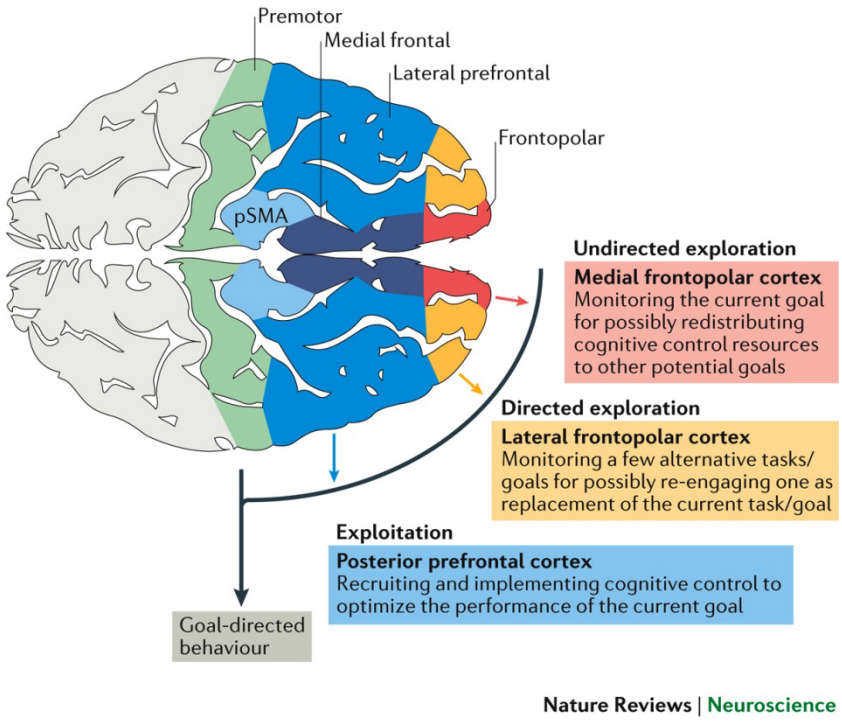

Figure 1-4. Functional model of the frontopolar cortex in humans. The medial frontopolar cortex is responsible for undirected exploration, which involves monitoring the current goal and potentially reallocating resources to other goals. In contrast, the lateral frontopolar cortex is responsible for directed exploration, simultaneously monitoring alternative goals with the potential for replacement. On the other hand, the posterior prefrontal cortex optimizes the performance of the current goal and primarily engages in exploitation. pSMA=pre-supplementary motor area. *Nature Reviews, Mansouri, F. A., Koechlin, E., Rosa, M. G. P., & Buckley, M. J., Neuroscience, Managing competing goals - a key role for the frontopolar cortex, 18(11), 2017, reproduced with permission from Springer Nature (SNCSC).*

Moreover, the neurotransmitter dopamine has been associated with exploration-exploitation decisions. Elevated levels of dopamine in the striatum were found to enhance exploratory behavior, as indicated by Verharen et al. (2019). On the other hand, decreased dopamine levels primarily dampened directed exploration compared to random exploration (Chakroun et al., 2020). These findings suggest that the effects of dopamine modulation may differ depending on the specific exploration strategy employed. In a foraging task, Le Heron et al. (2020) showed that dopamine specifically

modulates the influence of BRR but not FRR, while signaling the opportunity cost of rewards, i.e., the lost value of alternatives when an option is chosen.

Specifically, for foraging decisions, the aMCC was identified to play a key role in tracking the average value of the foraging environment (search value) and the costs of foraging (Kolling et al., 2012). Beyond that, opposing reward trends (average reward rate and recent reward rate), as well as expected future rewards have been associated with aMCC activity before the decision to leave (Wittmann et al., 2016). Contrarily, McGuire and Kable (2015) were not able to identify encoding of the value of quitting within the aMCC, rather, the vmPFC and the striatum reflected a dynamic reassessment of subjective value. While for most of the aforementioned evidence, functional magnetic resonance imaging (fMRI) was used, Hayden et al. (2011) was able to show that the cells in the ACC in macaques exhibited a rise-to-threshold signal in foraging decisions. Specifically, this study was the first attempt to demonstrate a temporal resolution of the foraging process in primates. The firing rate of neurons in this region gradually increased over time within a patch until the moment the monkey decided to leave. Specifically, neurons reached the firing threshold more quickly on trials where the monkey left the patch earlier.

In summary, some efforts have been made so far to find neural representations of foraging behavior and to parse out the role of dopamine in this process. However, there is an important gap in the literature because the foraging process is a "when decision" and thus highly temporally resolved (Kolling & O'Reilly, 2018). This process, which changes dynamically over time, cannot be imaged with sufficient resolution by fMRI. Therefore, it is necessary to use high temporal resolution methods like EEG in humans to map the exact neuronal processes in the course of the decision process.

1.1.5 Research questions and hypotheses

In conclusion, novel approaches try to illustrate more complex decision-making in humans, yet the evidence is sparse and there is a particular lack of studies investigating the temporal process of exploration-exploitation decisions. On a behavioral level, there are individual findings that people approach MVT in their behavior, albeit to certain constraints and biases. These studies show that participants tend to stay longer in

patches of higher yield than lower yield (FRR) and stay shorter in an overall rich environment than in a poor one (BRR). Nevertheless, these adaptations are often insufficient in magnitude to match the predictions of a normative model. There is still little evidence that examines human foraging behavior in comparison to a normative model and describes biases in this framework. There is also a need to replicate the limited evidence to date in different foraging paradigm contexts. This gives rise to several research questions for this thesis: How do people decide to leave a current option? How do people adapt to different contexts? How well do humans perform compared to an optimal MVT-inspired policy? To investigate these issues, we conducted two manipulations in a foraging paradigm with a continuous harvest phase. The manipulation of FRR involved altering reward decay across patches, while the manipulation of BRR involved varying the environment's quality through distinct patch probabilities and time costs during foraging. Based on the research questions and the literature reviewed above, I propose the following hypotheses:
In a foraging task with different reward environment, I expect individuals to adapt their decision to leave according to the change in FRR and BRR independently and thus mimic the behavior of an MVT-inspired optimal agent. As previous literature suggests, subjects will adapt to manipulations in the direction that MVT would predict, but not to a sufficient degree. Specifically, I hypothesize that participants will stay longer in patches with higher yields, while staying shorter in patches in a rich environment. While generically "overharvesting" (staying longer than optimal), participants will adjust their behavior better to a rich than to a poor environment.

On a neuronal level, very little is known about foraging-like decision processes so far. Although the role of the aMCC and of dopamine in the foraging context have been emphasized, there is a lack of temporally resolved imaging data on this decision process. The aim of the present work is to address this gap. Is there a specific neural mechanism underlying foraging decisions tracking different reward rates of foreground and background decisions? If such a mechanism exists, at what specific moments are these signals represented and how do they manifest? There are some options, how such a process might be represented in a neuronal signal. If, as the MVT suggests, a threshold must be reached to find the optimal time to leave, I might expect a signal representing an increase from initial foraging to reaching that threshold that determines patch leaving

decision. Another possibility is that the decision to leave the current patch is made fairly early in the foraging process, immediately after an initial phase, by participants extrapolating the decay function of the current patch. A third possibility is that individuals use simple heuristics that may lead to a signal associated with leaving immediately before the decision is executed (or at other times), without performing complex threshold calculations. While behavioral results can provide initial insights into addressing this question, a comprehensive and descriptive solution requires uncovering the underlying neural processes. Regarding the manifestation of the signals, there are some findings that associated activity in the theta frequency range in central regions with decision-making (Cortes et al., 2021; Jacobs et al., 2006), but there are no such findings yet in a specific foraging environment.

An intriguing finding by Cavanagh et al. (2012) revealed an increase in midfrontal theta activity in exploration-associated response-locked data, indicating that theta-band activity may signify the urge to reduce uncertainty and exercise strategic control during exploratory choices. Conversely, in exploitation-associated data, negative correlations were observed for medio/lateral/frontal theta and beta power. As a result, theta activity emerges as a potential candidate for representing a tracking signal in exploration-exploitation-related foraging decisions, although it should be noted that this hypothesis is highly exploratory in nature. Regarding the observed changes in beta activity, there is additional evidence linking beta signals to decision thresholds. Specifically, beta power lateralization reflects the state of evidence accumulation during decision formation, with its peak serving as a neural indicator of the decision threshold (Fischer et al., 2018; Kirschner et al., 2023; Rogge et al., 2022). As a result, it can be hypothesized that beta and theta band activity are associated with the decision to abandon the current patch in favor of a potentially more rewarding resource within a foraging paradigm.

1.2 Feedback processing

The decision-making process consists of multiple stages (***Figure 1-1***). The initial part of this thesis focuses on the valuation and choice phase within this cycle. In contrast, the objective of the second study is to explore how humans evaluate their actions and adapt to the consequences once a choice has been made.

Goal-directed behavior and performance monitoring encompass a range of processes that facilitate flexible adaptations. This can be visualized as a feedback loop, where the weighted differences between expected and actual action outcomes are utilized to initiate appropriate adjustments and improve outcome prediction (Ullsperger, Danielmeier, & Jocham, 2014). It involves comparing the anticipated outcome with the observed outcome by calculating prediction errors and modifying behavior to reduce these errors through a learning process. Consequently, reward learning plays a vital role in value-based decision-making, where feedback is employed through positive and negative reinforcement. Monitoring behavior can involve learning from deliberate errors that are internally processed or through external feedback. External feedback is particularly important to assess one's own performance when errors are not consciously recognized. The second study focuses on the processing of external feedback and the subsequent behavioral adaptations that ensue.

There are a lot of theories concerning how feedback is integrated in evaluation and learning. For a comprehensive overview, please see Ullsperger, Danielmeier, and Jocham (2014). In the following sections, I will focus on the most influential theories and their electrophysiological and behavioral evidence and describe open questions.

1.2.1 Theoretical and electrophysiological basics of feedback processing and performance monitoring

There are two main event-related potentials (ERP) associated with feedback processing: the feedback-related negativity (FRN) and the P300. The FRN is characterized by a negative deflection peaking around 200-300ms after feedback originating in frontocentral areas (Gehring & Willoughby, 2002; Miltner et al., 1997). Historically, the FRN has been assumed to be larger (i.e., more negative-going) for

negative compared to positive feedback (Miltner et al., 1997). The P300 (or P3, respectively) refers to a family of positive ERP deflections elicited by action-related stimuli around 300ms after presentation. There is the sharp and early P3a originating in more frontocentral areas responsible for attentional processes, while the more sustained P3b in parietal areas is associated with updating memory (Courchesne et al., 1977; Polich, 2007). Moreover, the P3b is associated with uninformative surprise coding (Donchin & Coles, 1998; Mars et al., 2008) in a way that it is larger (i.e., more positive-going) for unexpected or infrequent than expected/frequent events (Johnson & Donchin, 1980; Polich, 2007). Furthermore, the P3b seems to be involved in action-value updating (Ullsperger, 2017; Ullsperger, Fischer, et al., 2014) and acts as a bidirectional learning signal (Nassar et al., 2019). While the FRN seems to be responsible for early evaluation processes, the P3 reflects the transformation of these information into attentional, motivational, and working memory processes (Huvermann et al., 2021). This gives reason to believe that both ERP components represent different aspects of information processing.

One of the most important models in this area, which has been heavily debated over the past 20 years, is the reinforcement learning theory (RFL) of error and feedback processing proposed by Holroyd and Coles (2002). The model suggests that the FRN, like the error-related negativity (ERN), is an electrophysiological phenomenon of neural activity originating from the aMCC. According to this model, the FRN reflects a reward prediction error (RPE) that represents the discrepancy between an expected and an observed outcome. Thus, the FRN is supposed to scale with the size of the RPE being larger for unexpected outcomes (Holroyd & Krigolson, 2007; Weismüller & Bellebaum, 2016). Hajcak et al. (2005) investigated this phenomenon using a guessing task in which the expectation of the outcome was manipulated by changing the probability of winning. Contrary to the model's prediction, the FRN was not affected by the manipulation of expectancy, while the P3 increased as a function of unexpectedness. Since then, many studies have attempted to clarify what internal processes the FRN reflects (for a detailed overview, please see section 4.1). One of the proposed theories states that the FRN changes both as a function of valence and as a function of expectancy, representing a so-called signed prediction error (PE). A large FRN is elicited, when the feedback is worse than expected (negative RPE), whereas a smaller FRN is elicited, when the feedback is

better than expected (positive RPE; Fischer & Ullsperger, 2013; Hajcak et al., 2007; Walsh & Anderson, 2012). However, new approaches indicate that the FRN signal may in fact be a product of a superimposed positive-going deflection, the reward positivity (RewP), driven by better-than-expected outcomes (Baker & Holroyd, 2011; Krigolson, 2018; Proudfit, 2015). Recent research suggests that these are, however, independent processes corresponding to the representation of the positive and negative prediction error (Bernat et al., 2015; Cavanagh, 2015; Hoy et al., 2021; Zheng & Mei, 2023). In this context, it is important to highlight that the interpretation of results relies on how the FRN is defined and quantified. Different approaches exist, such as quantifying the FRN as the difference wave between losses and wins or as a distinct component resembling the N2 wave following feedback. The use of different paradigms and methods to represent the FRN makes results difficult to generalize. There continues to be a lack of studies that attempt to resolve this confound.

Regarding the P3 component, there is supporting evidence that indicates its association with unexpectedness (Fischer & Ullsperger, 2013; Hajcak et al., 2007; Walentowska et al., 2016). However, when considering the effects of valence, the results are still mixed and unclear (Severo et al., 2018; Yeung & Sanfey, 2004).

1.2.2 Adaptive behavior after feedback

To effectively monitor our performance, it is crucial to continuously evaluate feedback on our actions and, if necessary, respond by adjusting our behavior. Several phenomena can indicate behavioral adjustments following an error, including post-error slowing (PES; Debener et al., 2005), post-error reduction of interference (PERI; King et al., 2010), and post-error improvements in accuracy (PIA; Marco-Pallarés et al., 2008). For a comprehensive review, please refer to Danielmeier and Ullsperger (2011). However, these phenomena are typically observed in the context of aware errors that are promptly corrected. The evidence becomes less clear when considering unaware errors that require external feedback for detection. Several studies suggest that activity in the anterior midcingulate cortex (aMCC) is associated with behavioral adjustments after errors, serving as a mediator for translating changes in action outcomes into the necessity for adaptive behavior (Danielmeier et al., 2011; Ullsperger, Danielmeier, & Jocham, 2014). Behavioral adaptations appear to depend on the valence of feedback and

its interaction with expectedness, with corresponding correlations to the amplitude of the FRN. Specifically, larger behavioral adaptations were shown after unexpected negative feedback (Holroyd & Krigolson, 2007). Recent evidence also indicates that the amplitude of the RewP predicts timing behavior for subsequent trials (Yan et al., 2023). Despite this progress, our understanding of the processing and execution of adaptive behavior after feedback, as well as the underlying neuronal mechanisms, remains limited.

1.2.3 Research questions and hypotheses

Despite extensive research in this area, numerous unanswered questions remain. There are conflicting findings as to whether, and if so, what types of RPE the FRN represents. Additionally, discrepancies in the quantification of FRN/RewP as an ERP component have yielded non-comparable and divergent results. It is crucial to adopt a quantification-independent approach that accurately represents post-feedback neuronal processes while independently capturing the influences of valence and expectedness on the neuronal signal. The P3 component appears to reflect unexpectedness, but its relationship with the valence of feedback is still ambiguous and requires further clarification. Scant findings exist regarding adaptive behavior after feedback, necessitating additional research that explores the association between feedback-processing FRN and P3 signals and behavioral adaptations. In light of these gaps, this thesis aims to address the existing issues by employing a novel methodology that can describe the influences of external factors on the neuronal signal without relying on specific ERP quantification. This will be achieved through a single-trial regression approach. The research questions to be addressed are as follows: Which type of RPE is reflected in the neuronal signal during the FRN period - signed or unsigned? Does the neuronal signal during the P3 timeframe solely reflect unexpectedness, or does it also incorporate the valence of feedback? Is the neural signal during the FRN and P3 periods associated with behavioral adaptations to feedback? Based on previous findings, the following hypotheses can be formulated: Firstly, an interaction effect of expectedness and valence on the neural signal within the FRN timeframe is expected, supporting the notion of a signed RPE. Furthermore, I anticipate a significant impact of unexpectedness on the neuronal signal during the P3 latency. Additionally, it is expected that the neuronal signal within the FRN

latency correlates with behavioral adjustments after feedback, with larger adjustments and increased neuronal activity following negative feedback.

2 General methodology

The following section provides a short overview and description of the key experimental and analysis methods that were used in the present dissertation. A more detailed description can be found in the methods sections of the empirical studies in section 3 and 4.

2.1 Experimental paradigms

In a broader sense, foraging consists of searching, locating, and collecting resources across multiple patches with the main goal of maximizing the overall rate of resource intake within a given time. As a subcategory, patch-leaving paradigms focus specifically on the decision of when to leave a resource patch, considering factors related to resource depletion and opportunity costs. The goal of developing the paradigm used in the first experiment was to map the entire decision-making process as described in ***Figure 1-1***. That's why, a binary value-based decision was combined with a patch-leaving decision. The present work focuses on the part of this process that involves patch-leaving decisions.

Currently, there is a lack of established foraging paradigms. Constantino and Daw (2015) employed a paradigm in which participants harvested trees and made distinct choices regarding whether to continue harvesting a particular tree or travel to another one. However, the present study focused on investigating the *continuous* foraging process using EEG. To achieve this, we implemented a continuous patch-leaving phase. In this phase, participants determined their foraging duration on a patch based on the desired reward threshold, as the reward continuously diminished over time. Moreover, environmental differences can be manipulated by varying the average patch quality and travel time between environments (Gabay & Apps, 2021). In our study, we presented these variations in a block-wise manner with fixed time periods for participants. This approach allowed us to effectively utilize EEG as a research method and address our research questions.

In study II, to specifically generate certain EEG signals, we employed an established task called the time estimation task. In this task, participants were required to estimate

the duration of one second by pressing a key, and they received feedback indicating whether their estimation was accurate (positive feedback) or not (negative feedback). Outcome expectancy was manipulated by adjusting the target time window around one second, either increasing or decreasing, within which participants' responses were considered correct. For instance, if a response was marked as incorrect in an environment with a wide target time window (making it easier to hit), it was regarded as an unexpected outcome. It's important to note that learning in this task was limited since the feedback provided did not specify the direction in which participants should adjust their behavior to provide accurate estimations. This distinction sets it apart from probabilistic learning paradigms commonly utilized in investigating feedback processing.

2.2 Computational modeling of behavioral data

As mentioned earlier, computational modeling can greatly enhance decision research by creating models that reveal the underlying computational processes and latent variables influencing decision-making behavior. Ideally, one can then identify a neural signal that represents these decision variables.

Previous findings (Harhen & Bornstein, 2023; Kolling & Akam, 2017; Steixner-Kumar & Gläscher, 2020; Wittmann et al., 2016) indicated that an optimal fit to participants' behavior is achieved through a combination of model-based and model-free approaches. In this context, the term "model-based" typically denotes a sophisticated reinforcement learning approach leveraging prediction errors for anticipating future experiences. Conversely, the "model-free" approach is often linked to a straightforward MVT-like learning rule featuring an exit threshold. However, recent perspectives challenge the oversimplified categorization of model-based and model-free approaches (Collins & Cockburn, 2020).

In this context, it is necessary to distinguish between normative approaches and models that fit parameters to an individual's behavior. The latter helps to depict individual variations, biases, etc. via these parameters. Normative approaches involve mathematically solving decision problems optimally, such as the MVT within the foraging context. Utilizing a normative model allows us to explore deviations from optimal decision strategies and their mechanisms. Therefore, the main focus in study I was to

develop an optimal agent inspired by the MVT framework. This agent utilized each participant's specific environment to determine the ideal patch-leaving solution. On the one hand, this model demonstrates the generalizability of a continuous patch-leaving paradigm to the MVT approach. On the other hand, this individual optimum served as a reference for comparing the observed behavior of participants and describing deviations from the optimal strategy. Subsequently, these deviations were associated with neuronal signals extracted from EEG data in order to explore the neural correlates of behavior.

2.3 Regression models for EEG data

As previously mentioned, the high temporal resolution of EEG makes it well-suited for studying continuous foraging processes and feedback processing. To disentangle various influences on the neuronal signal, I employed a robust single-trial regression approach (Fischer et al., 2016; Fischer & Ullsperger, 2013) to analyze the EEG data. Single-trial EEG activity at each electrode and time point was regressed against experimental and behavioral parameters. The models, which are little susceptible to outliers, were employed in the time- and the time-frequency domain, first within and then across subjects in the following linear equation (*(2-1*):

$$Y = intercept + \beta_1 Reg_1 + \beta_2 Reg_2 + \ldots + error \qquad (2\text{-}1)$$

This mass univariate approach yields individual b values for each electrode and time point for each subject. To ensure comparability of predictors within and between subjects and to address multicollinearity, the b values were standardized by their standard deviations before averaging across subjects. This approach allows to investigate simultaneous influences of multiple independent variables while preserving the high temporal resolution of the EEG.

3 Study I: Temporal EEG dynamics of foraging decisions in humans

3.1 Introduction

During the search for a new job, individuals must consider various factors: comparing alternative options to the current job, contemplating potential future offers, and determining the opportune moment to change jobs depending on labor market conditions. This challenging decision-making process in volatile environments can be described by the term "Foraging decision". This type of decision-making involves deciding whether and for how long to continue exploiting a rewarding patch or to explore new opportunities – a challenge referred to as the exploitation-exploration dilemma. Humans face a trade-off between the benefits of acquiring reward and the costs associated with foraging, such as time and energy expended, as well as potential risks. As resources become depleted, it becomes less and less efficient to continue exploiting the current patch, and the costs of moving to a new patch may become outweighed by the benefits of finding a new resource. The marginal value theorem (MVT; first proposed by Charnov, 1976) provides a normative framework for understanding how humans achieve this balance and find the ideal leaving time: one should stay in a patch of resources until the rate of energy gain (i.e., the amount of resources obtained per unit time) in the current patch falls below the average rate of energy gain across all available patches in a specific environment. The former is also often described as foreground reward rate (FRR), the latter as background reward rate (BRR). This theory suggests that a longer stay in richer patches is beneficial compared to poorer patches, while a shorter stay on patches in an overall rich environment consisting of multiple rich patches is beneficial compared to a poor environment. This phenomenon has already been demonstrated in several studies in humans (Constantino & Daw, 2015; Le Heron et al., 2020; Wolfe, 2013) and primates (Hayden et al., 2011), sometimes with restrictions (Garrett & Daw, 2020; Wittmann et al., 2016; Wolfe, 2013). Recent evidence examines the "overharvesting" (staying too long) of participants that is sometimes observed and reveals that MVT is insufficient to explain behavior because it does not account for uncertainty as a function of task complexity (Harhen & Bornstein, 2023). In addition, one study showed that participants do not adequately adjust their expectations when the

environment worsens, leading to suboptimal decisions (Garrett & Daw, 2020). On a neurobiological level, the average value of the foraging environment, the cost of foraging, and expected future rewards were associated with the aMCC (Hayden et al., 2011; Kaiser et al., 2021; Kennerley et al., 2006; Kolling et al., 2012; Kolling et al., 2016; Wittmann et al., 2016). Others (McGuire & Kable, 2015) additionally showed an involvement of the ventromedial prefrontal cortex (vmPFC) in the foraging process, as it is associated with the comparison of the values of different options. Moreover, Le Heron et al. (2020) found that dopamine specifically modulated how sensitive people were to the richness of the environment.

While evidence is growing that human foraging decisions can be well described, at least at a conceptual level, by previous models such as the MVT with certain additions, many questions still remain open. It is unclear in what way which information is used by subjects to make an informed foraging decision. Moreover, neurobiological insights into this process are still sparse, and so far, fMRI has been used significantly to draw conclusions about localization. As a consequence, there is yet little knowledge about the timing of neurological dynamics of this process, although it is a "when decision". Thus, an important component for understanding foraging decisions is missing at this point. EEG is a highly suitable approach for addressing this research question but has not been used to investigate foraging decisions in patch-leaving paradigms so far. To bridge this gap, the purpose of this paper is to explore the temporal component by introducing a novel foraging paradigm coupled with concurrent EEG recording. At the behavioral level, we expect participants to approximate in their behavior an ideal agent as proposed by MVT, in terms of close-to-optimal adaptation to changes in FRR and BRR. Specifically, we anticipate that participants spend more time on patches with higher quality and to prolong their stay on patches within a poor environment. Building on previous evidence, we generally predict overharvesting and poorer adaptation in poor compared to rich environments.

At the neurobiological level, we expect to see foraging-related dynamics of neural activity that track changes in FFR and BRR. Such decision signals could initially occur at the beginning of a harvest phase on a particular patch and already prime the ideal leaving time. Alternatively, such a signal could also occur as a persistent tracking signal leading up to the decision, possibly indicating some sort of upramping to a calculated threshold

that needs to be crossed. Hayden et al. (2011) provides initial evidence for this in a foraging study with macaques: neurons firing in the primate dACC predicted when monkeys left a patch, and these responses increased with the time spent in the current patch until neural responses reached a threshold. It is still unknown, how a tracking signal might look like in humans. Because midfrontal theta has been associated with cognitive control (Cavanagh & Frank, 2014) and strategic control during exploratory choices (Cavanagh et al., 2012), while previous evidence has linked foraging decisions to the aMCC, theta tracking signals in frontal areas could represent such a marker. Moreover, pre-response reduction in beta power over the contralateral motor cortex has been demonstrated to reflect decision-related variables (Donner et al., 2009; Fischer et al., 2018; Pape & Siegel, 2016). Activity in the beta band can serve as a proxy for response thresholds (Kirschner et al., 2023) and signifies integration of decision evidence (Rogge et al., 2022). Hence, a decrease in beta activity not only signals motor action preparation, rather it provides a readout of action selection. In conclusion, we expect activity in the theta band and a reduction in beta power to encode the decision to leave in patch-leaving environments.

3.2 Methods

3.2.1 Participants

N = 137 healthy participants were recruited at the Otto-von-Guericke University Magdeburg (11/2020-04/2022). Exclusion criteria were any present or past psychiatric or neurological disorders, drug abuse, alcohol intake at day of study, and use of neurological-acting medication. 3 participants had to be excluded due to an interruption of task execution, therefore these subjects did not complete all conditions equally. The final sample for the analysis of behavioral data consists of 81 female, 52 male and 1 diverse participants between 18-40 years (*M* = 23.37; *SD* = 4.99). The study was approved by the Institutional Review Board/Ethics Committee the University of Leipzig (285-09-141209) and written informed consent was obtained from all participants after briefing prior to study enrolment. The study was conducted in accordance with the Declaration of Helsinki. For their participation, the subjects received a payment of 8 euros per hour or credit points for their studies. In addition, all participants were able to earn a bonus of 0.02 euros per point won in the paradigm.

3.2.2 Experimental paradigm

We have developed a novel patch-leaving task (***Figure 3-1***) where the main goal of the participants is to gain as much reward (gold) as possible within a given time. Each trial starts with a value-based binary choice between two options (islands). The objective value of the options is defined by the possible gain of the option (green bar), the possible loss (red bar), and the probability of loss (number of pirate symbols above: 1=25%; 2=50%; 3=75%). In addition, information is given about the accumulated reward won (blue bar at the top; in points) and the time remaining for foraging (sun; in days). Participants decide which option to select by pressing a button with their left or right thumb. This is followed by a varying travel time to the chosen option (continuously filling blue bar), with time lost to foraging. During the harvesting phase, a filling green circle is displayed, which represents the reward collected by the participants. The filling circle is based on 9 different decay functions (PatchScaler; ***Figure 3-1 Bi and Bii***), so that the circle fills up more and more slowly over time. By pressing the button with the right thumb, the participants decide to leave the option (patch) to search for more reward elsewhere, and thus the harvesting phase is terminated. At the end of the trial,

participants receive feedback on the reward they collected during the harvesting phase and, in a second step, on whether and how much of that reward they lost again (due to a pirate attack). The feedback was displayed for 7 sec on average. Throughout the trial, except during the binary decision, time is lost to foraging.

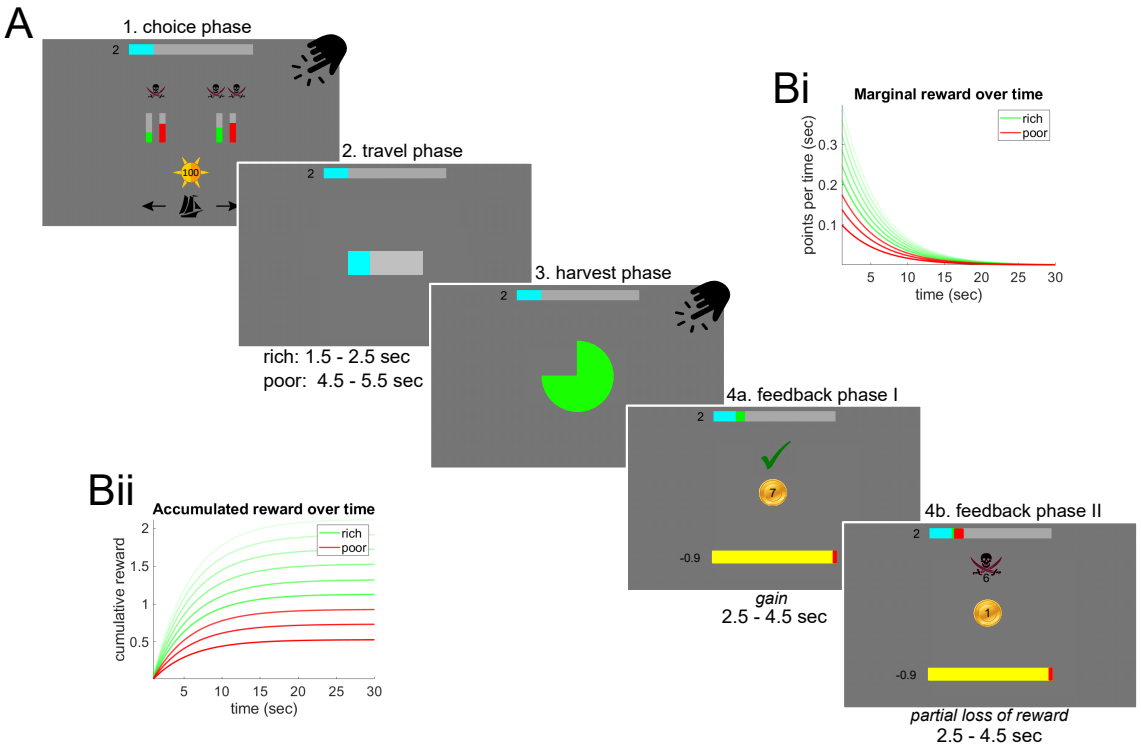

Figure 3-1. Task Design. 137 healthy volunteers had 30min (illustrated by the sun and the yellow bar; 100 days) to harvest as much reward as possible while EEG (64 equidistant electrodes, sampling rate of 500Hz) was recorded (A). They started with a value-based choice between two options varying in value (1), had to wait in a travel phase (2), decide how long to stay and harvest the chosen patch (filling circle, 3) and received feedback at the end about the reward gained (4a) and what was left of it after a possible loss (4b). For 15 min, the environment was rich with more valuable options to choose from and a shorter varying travel time, while for the other 15 min, the environment was poor. The patches were defined by different decay functions (Bi). Bii shows the accumulated reward of the patches. The more transparent the line, the higher the quality of the patch. The rich environment consisted of the 6 patches with the highest quality (green), in the poor environment all 9 patches could occur (red and green).

A total 30min (100 days) is available, including 15min for each of two conditions. Participants are presented with a rich (summer) and a poor (winter) environment, with options with higher objective value (gain – loss*probability) available in the rich environment and shorter average travel time (*M* = 2 sec) than in the poor environment (*M* = 5 sec). Participants went through each condition once (counterbalanced over participants) knowing which environment they are in by the instruction. Due to task design, in the rich environment occurred only the 6 highest quality patches (***Figure 3-1***

Bi; Bii green), while in the poor environment, all of the 9 patches, with higher probability for low quality patches (***Figure 3-1 Bi and Bii red and green***), could occur. Based on a collection of predefined trials, specific trials were randomly drawn and presented until the time to forage expired, so that each participant experienced an individual environment (mean number of trials = 116; *SD* = 11). Before the experiment, participants undergo a training session to ensure they understand the complex instructions and objective of the task.

In order to disentangle neural correlates uniquely attributable to cognitive foraging processes from motor-related processes, we added a control paradigm very similar to the main task to compare the results of both versions with respect to their different demands at the end of each session: while the patch-leaving task requires intensive top-down involvement to determine the optimal time to leave in a given environment, the motor task was very simple and mainly requires a motor response at a specific time. It consists of 50 trials randomly selected from the previous task (10% of the longest leaving times, 10% of the shortest, rest randomly), in which participants are presented with filling green circle indicating their previous leaving time with a black marker. Their task is to press the button when the circle reaches the marker. All preprocessing and analysis steps are congruent with the main task. Results can be found in the supplement information.

3.2.3 EEG acquisition and preprocessing

Electroencephalic signals were continuously recorded at 500 Hz sampling rate with BrainAmp MR plus amplifiers (Brain Products) from 64 Ag/AgCl sintered electrodes, which were mounted in an elastic cap according to the extended 10-20 system with impedances kept below 5kΩ. The ground electrode was placed at the sternum. Electrodes to capture horizontal and vertical eye movements were mounted next to both eyes and above and below the left eye. The signal was online referenced to the left mastoid. The recorded data was high (0.3Hz) - and low (42Hz) -pass filtered and re-referenced to common average. In the present study, we focus on the harvest phase. Response-locked data (decision to leave the patch) was epoched from -3000 to 500ms after response, whereas stimulus-locked data (start of the green circle) was epoched from -500 to 3000ms after stimulus onset. Artifacts were excluded based on an

automated algorithm that rejected at least 5 epochs and no more than 10% of the trials, while the rejection criterion for signal outliers was a deviation of more than 5 SD from the mean probability distribution of the EEG signal (Delorme et al., 2007) and could be adaptively adjusted to meet the previously mentioned criteria. In addition, epochs were demeaned and submitted to adaptive mixture independent component analysis (AMICA, Palmer et al., 2012). We used a combination of automated procedures (Corrmap-approach, Viola et al., 2009) and the rating of two independent raters to exclude artifactual components from the data. For the analyses of the behavioral and EEG data, we calculated multiple robust regressions (Fischer et al., 2016; Fischer & Ullsperger, 2013). A baseline of -300 until 0ms prior to stimulus onset was used and subtracted from the signal. EEG datasets for which ICA did not converge or too few trials for specific regressors existed were excluded (3 subjects), resulting in a final EEG sample of 131 participants. For EEG and behavioral analysis, EEGLab v2021.0 toolbox (Delorme & Makeig, 2004) and customized code written in MATLAB R2019b version 9.7.0.1586710 Update 8 (MathWorks) was used.

3.2.4 Data analysis

3.2.4.1 Analyses of behavioral data

As mentioned before, we focused on the harvest phase of the paradigm, because we were interested in patch-leaving decisions. The instantaneous reward rate at leave time (*iRR*) served as the outcome variable, as we were interested in the influence of several task-specific characteristics on the time-reward ratio at which participants exit the current patch. This parameter was calculated by subtracting the accumulated reward at the timepoint before the decision to leave from the accumulated reward at the time of the decision to leave, relative to the time resolution. Highly correlated with this outcome measure is patch leaving time (*PLT*), which describes the time spent on the current patch. Several regressors of interest complemented the regression analyses: the type of environment (rich/poor; *Environment*); quality of the patch (9 different decay functions of reward development; *PatchScaler*); pirate raid/loss in previous trial (y/n; *Pirates*); the time left in the current environment (*DaysLeftB*); the average of the amount of the last 5 rewards participants have received (*Last5*); the duration of the travel time

prior to the harvest phase relative to the mean travel time of the other environment (*relSailing*). The overall model was defined as:

$$iRR = \beta_0 + \beta_1 \times Environment + \beta_2 \times Pirates + \beta_3 \times PatchScaler + \beta_4 \times DaysLeftB + \beta_5 \times Last5 + \beta_6 \times relSailing + \varepsilon \quad (3\text{-}1)$$

The type of environment was represented by the BRR, whereas the FRR was operationalized as PatchScaler. In the context of MVT, an ideal agent would spend less time on each patch if the background reward rate (rich environment) is high, while a high foreground reward rate should result in an increase in time spent on the patch. Therefore, the iRR is supposed to be stable across different PatchScalers, however the iRR should be higher in rich environments for optimal success. We expect no or minor influence of other regressors onto the iRR. To quantify participants' performance on the task, we used a measure of negative efficiency by comparing participants average reward rate of leave to an ideal agent and calculating the percentage deviation from the optimum. We linked this measure to the EEG signal during the harvest phase, described in more detail in 3.2.4.3.

All parameters were normalized, and reaction times were log-scaled. To compare participants' behavior with an ideal policy, we simulated an optimal agent, which is described in the following section.

3.2.4.2 Computational Modeling

We developed a model that mimics each participant's individual environment while maximizing rewards under time pressure. The optimal agent used participants' choices in the binary choice phase to determine the optimal time to stay on a particular patch in a given environment. According to MVT, the optimal leave time on each patch was defined as the time point when the change of reward over time (FRR) exceeds the average reward rate (BRR). Separately for each environment and based on individual participants' environment and binary choices (*cp*), the constant average reward rate (*en*) was fitted in an iterative process until it reached a negative local minimum indicative of the maximum reward per unit time. The optimization was executed employing the Matlab function "patternsearch," guided by Eq. (*3-2*). Within this equation, the variable t_{leave} symbolized the ideal leave time based on the exponential function *f* (Eq. *3-3*). This function was randomly selected from a predefined set *k* of options contingent on the

environment (*env,* Eq. *3-4*). *N* represented the trial count, *TT* stood for the averaged travel time depending on the environment, *loss* denoted possible loss, and *lp* indicated loss probability. It follows that the total payout per unit time (*Loss*) was defined as the cumulative amount of reward earned upon leaving at the time of an MVT-derived threshold, minus the potential loss of the chosen option, divided by the time required to obtain this reward. Subsequently, the fitted average reward rate corresponding to the maximum payout was employed to solve the rich/poor environment, resulting in optimal leaving times/ iRR for the nine patches (six in poor environment) in each participant's unique environment. The results were used to confirm the assumptions of the MVT in the context of our task and to identify participant deviations from ideal behavior when compared to empirical data.

$$\forall t_{leave}: f_k'(t_{leave}, env) = en \quad Loss_{env} = -\frac{\sum_{n=1}^{N}((f_k(t_{leave},env) - loss_k \cdot lp_k)\cdot cp_n)}{N \cdot TT_{env} + \sum_{n=1}^{N}(t_{leave} \cdot cp_n)} \tag{3-2}$$

$$f_k(t_{leave}, env) = a_{env} \cdot e^{\frac{-t_{leave}+1}{5}} \tag{3-3}$$

$$k \sim U(env) \tag{3-4}$$

$$a_{rich} \in \{4.99, 5.83, 6.75, 7.63, 8.48, 9.36\}$$

$$a_{poor} \in \{2.34, 3.25\ , 4.11, 4.99, 5.83, 6.75, 7.63, 8.48, 9.36\}$$

3.2.4.3 Analyses of EEG data

For the raw time-frequency decomposition, we analyzed 80 frequencies between 1 and 30 Hz with the help of Morlet wavelet cycles between 4 and 10. A baseline of -300ms to stimulus onset was used for the harvest-onset-locked data, while -2800 to -2600ms before response served as baseline for response-locked data. Results are converted to changes in decibels from baseline.

We employed multiple robust regression analyses within (1st level) and across subjects (2nd level) in both the time domain and time-frequency space (see Eq. *3-1*; Fischer et al., 2016; Fischer & Ullsperger, 2013) with the same parameters as described above. Regression was applied to -2500 to 300ms relative to foraging response, whereas regression of stimulus-locked data was applied -300 to 2500ms relative to harvest stimulus onset. To eliminate the possibility of other events confounding the analyses, we included only trials with a minimum time of 2500ms remaining on the patch (mean trials left = 101; *SD* = 10). For the time-frequency decomposition within the regression,

complex Morlet wavelets between 1 and 30 Hz in 20 logarithmically steps were used. The number of wavelet cycles was 4.5 and changed as a function of frequency. This resulted in maps of beta weights at each electrode and time point (and frequency) across participants.

Beta weight maps were thresholded (threshold for clustering: p=.05, one tailed). Clusters were identified as temporally contiguous signals that shared a common effect sign and exceeded the threshold. Cluster mass was calculated as the average absolute beta weight within a cluster times its size (number of timepoints or timepoints x frequencies contained in the cluster). To correct for multiple comparisons, the cluster mass for each cluster was compared to a permutation distribution generated by iteratively flipping the sign (1000 permutations for time-domain data; 10000 for time-frequency data; Nichols & Holmes, 2002).

$$EEG\ amplitude\ /\ power\ =\ \beta_0\ +\ \beta_1 \times\ Environment\ +\ \beta_2 \times Pirates\ +\ \beta_3 \times\ PatchScaler\ +\ \beta_4 \times DaysLeftB\ +\ \beta_5 \times Last5\ +\ \beta_6 \times relSailing\ +\ \varepsilon \quad (3\text{-}5)$$

To calculate beta activity, the time-frequency decomposition was computed for different trial clusters according to the corresponding PatchScaler, depending on how much reward they provide, and averaged across participants. Subsequently, the time-frequency decomposed data from frequencies 13 to 30 Hz were extracted and averaged. To show beta activity for differentially efficient participants, the extracted beta signals were averaged across performers with good and poor efficiency, which were determined by a median split of the participant group based on their efficiency.

3.3 Results

3.3.1 Behavioral Results

Results of the overall behavioral regression (Eq. *3-1*) show that the PatchScaler has the greatest influence on iRR (***Figure 3-2 A***, mean beta = 0.42, $t(133) = 15.32$, $p < .001$ (bonf cor), 95% CI [0.37, 0.48]). The better the current patch is (or the more reward it yields over time) that participants are on, the higher the iRR at the time of leaving. The environment where participants are located seems to play a less important role, yet participants tend to leave at a higher iRR in the rich environment compared to the poor environment (mean beta = -0.27, $t(133) = -3.13$, $p < .05$ (bonf cor), 95% CI [-0.49, -0.04]). ***Figure 3-2 B*** shows the distribution of negative efficiency as a measure of performance representing the deviation from an optimal agent for each participant in percent. The majority of subjects deviate from the ideal by only about 6%, with some deviating by as much as about 27%.

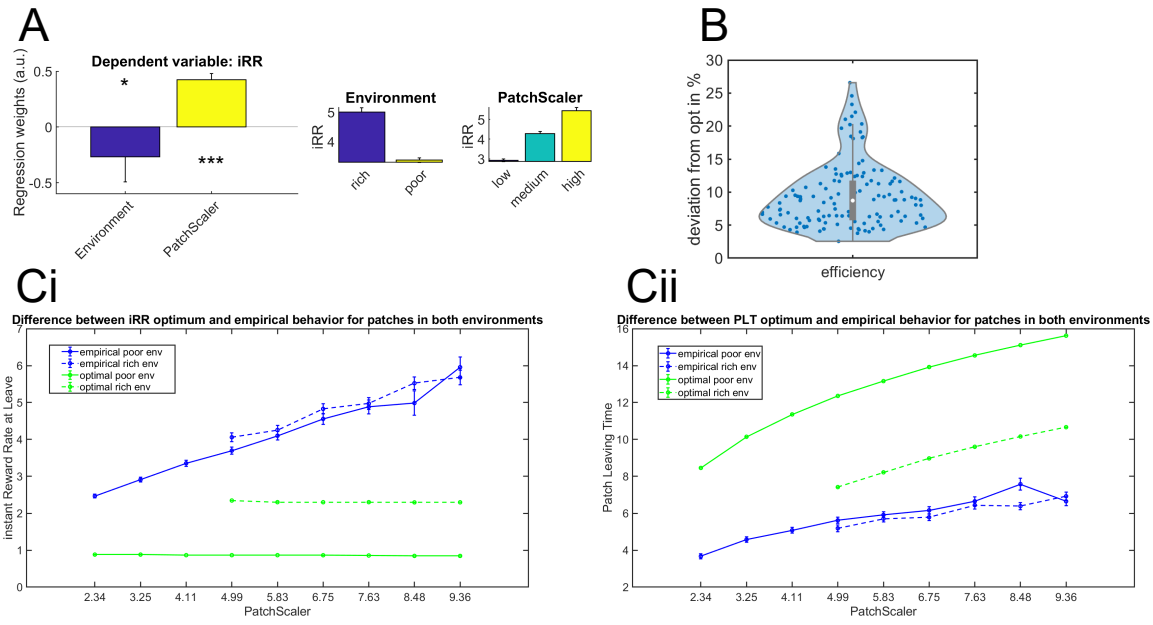

Figure 3-2. The effect of Environment and PatchScaler onto the instantaneous reward rate at leave time (iRR) and comparison of empirical with model data. Environment and PatchScaler have a significant influence onto iRR (**A**). Most participants deviate in their efficiency by about 5-10% from the optimum (**B**). Participants adjust the iRR depending on PatchScaler, while the optimal agent would propose a stable iRR that only differs between environments (**Ci**). Participants adapt to the patch leaving time (PLT; blue), but not to a sufficient degree as would be predicted be the agent (green, **Cii**).

Figure 3-2 Ci and Cii show the averaged data of an agent behaving ideally in the individual environments of the participants (green). As the MVT would predict, the iRR should remain stable between different PatchScalers, while it should be higher in rich environments than in poor ones. In contrast, the more reward the patch yields (PatchScaler), the longer the PLT should be, and it should be higher in poor environments than in rich ones. In comparison, participants actually show higher PLTs the higher the PatchScaler (blue), but not to the extent predicted from the normative model. At the same time, this means they leave the patch at suboptimally higher iRRs the higher the PatchScaler. Participants show a tendency to adapt to the environment in an optimal direction, but not to the sufficient extent that the agent would suggest. Overall, participants spend too little time on patches.

3.3.2 Electrophysiological Results

3.3.2.1 Stimulus-locked data

Results of robust single-trial regression on the EEG signal around the harvest onset reveal two periods where PatchScaler has a significant influence, namely a wide range from about 200ms to 1200ms after stimulus onset at frontocentral sites and a later period around 2400ms at centro-parietal sites (***Figure 3-3 Ai; Bi***). The sign of the beta weights reverses within the first period, so that the EEG signal is initially more positive the better the patch, and from 1000ms after the onset of harvest, the signal is more negative the better the patch quality. Starting at 2400ms, there is a positive shift in the signal that becomes more pronounced with improved patch quality again. The environment also influences the EEG signal in a more sustained fashion: from around 500ms on, there is an increase in the EEG signal with a more positive signal in the poor environment compared to the rich one (***Figure 3-4 Ai; Bi***).

We computed the same regression analysis as above using the time-frequency decomposed data. After cluster-based correction of the results (***Figure 3-3 Di***), we find a significant influence of PatchScaler in the form of a decrease in the frequency range of delta and theta starting 1500ms after stimulus onset in fronto-central regions. The better the quality of the patch becomes, the less of delta/theta activity occurs. Additionally, a decrease in the frequency range of beta differentiates between patches in more parietal areas from 600ms to 1600ms after stimulus onset.

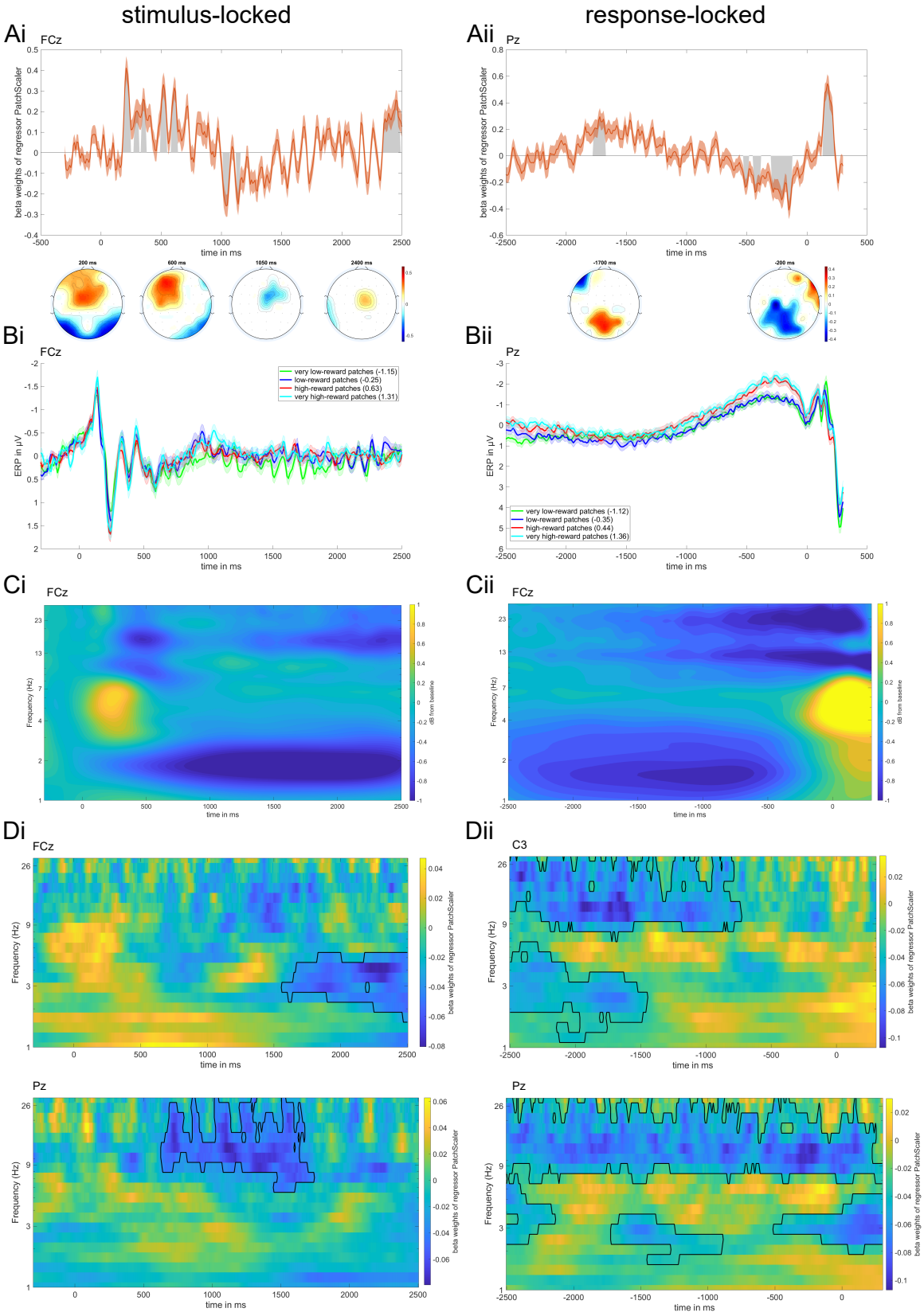

Figure 3-3. Effect of PatchScaler onto the EEG signal and raw time-frequency decomposed data. The left part of the figure shows the stimulus-onset locked results, while

(**Figure 3-3 continued)** the right part shows the response-locked results of the regressor PatchScaler. While the signal is more positive the better the patch quality already in the initial harvest phase in more frontal areas, this effect reverses just before the decision to leave is executed (**Ai, Aii, Bi, Bii**). **Ci** and **Cii** show the raw time-frequency decomposed data of the stimulus-locked and response-locked time window. There is an initial increase in theta power followed by a broad decrease in delta and beta frequency range until the decision is made. Results of the regression in the time-frequency space show decreases in the frequency range of beta, theta, and delta depending on PatchScaler that are most pronounced in the period before the decision to leave is executed (**Di, Dii**). Gray areas and black contours show significant clusters after permutation-based correction for multiple comparisons.

Due to environmental changes, there is a decrease in the beta frequency range for the maintained effect of the environment as well, while an increase in the frequency range of delta occurs right after stimulus onset in midfrontal areas. This is followed by a broad decrease in the theta frequency range (***Figure 3-4 Ci***).

3.3.2.2 Response-locked data

Results of robust single-trial regression on the EEG signal around the response when leaving the patch reveal two periods before the button press in which PatchScaler has a significant influence, namely earlier around -1800ms and immediately before button press from -500ms (***Figure 3-3 Aii; Bii***). In the first period, the beta weights have a positive sign, assuming a more positive EEG signal the better the patch, whereas the period immediately before leave decision suggests a more negative signal for better patches. Unexpectedly, both effects occur in more parietal areas. Again, the environment has a sustained influence onto the EEG signal with a more positive signal for the poor environment compared to the rich (***Figure 3-4 Aii; Bii***).

In terms of the time-frequency decomposed data, PatchScaler causes a broad decrease in beta frequency range before button press in midfrontal and parietal areas (***Figure 3-3 Cii; Dii***). Additionally, there are decreases in the frequency range of delta and theta that start in the early phase and last until the button is pressed. The effects mentioned before are most pronounced in fronto-central to parietal regions and contralateral to motion. Furthermore, the environment affects the EEG signal over parietal areas from around 2000ms to 1000ms before the decision to leave in the form of a decrease in the delta frequency range (***Figure 3-4 Cii***).

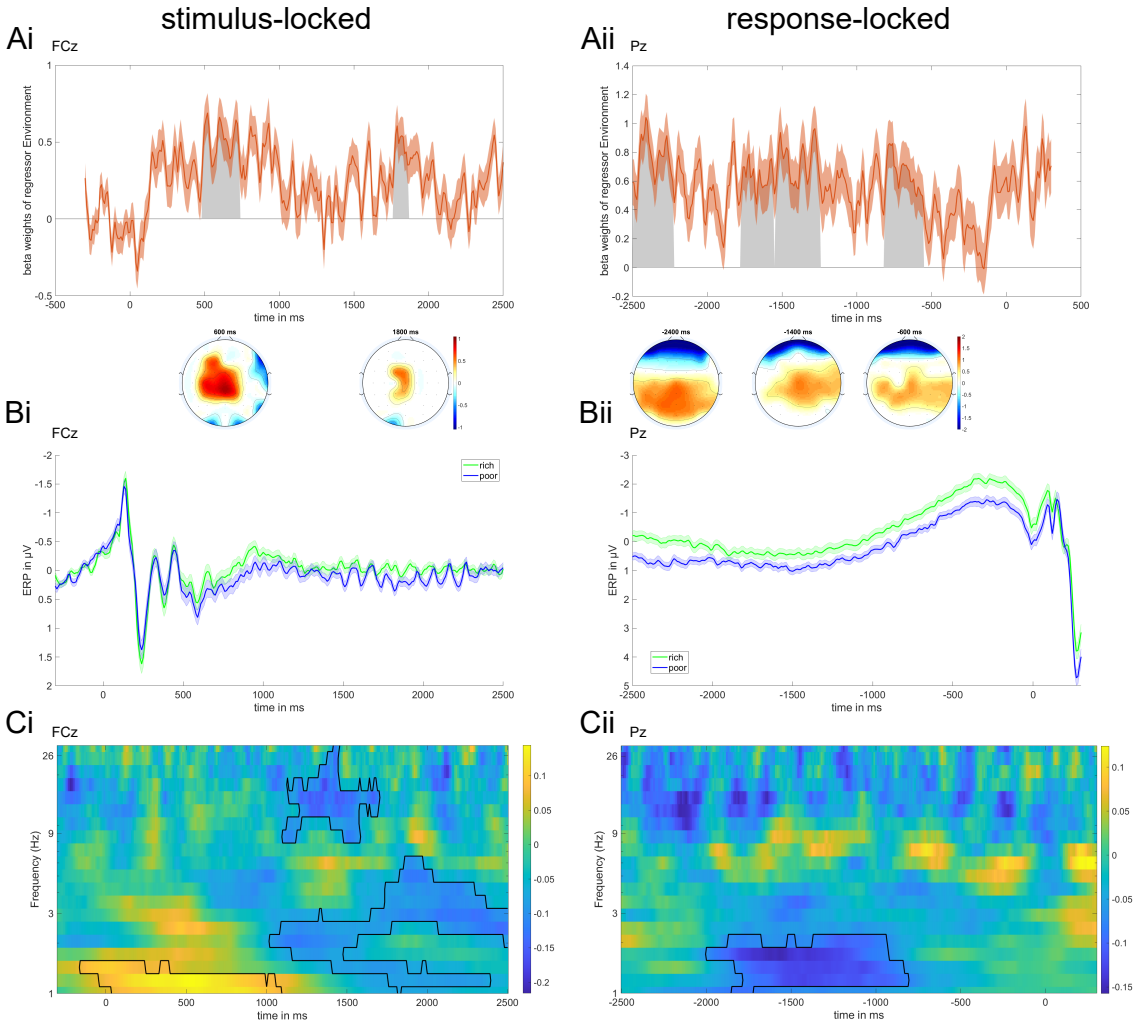

Figure 3-4. Effect of Environment onto the EEG signal. The left part of the figure shows the stimulus-onset locked results, while the right part shows the response-locked results of the regressor Environment. The signal is more positive in a poor compared to a rich environment and this effect is most pronounced in the period before the decision to leave in parietal areas (**Ai, Aii, Bi, Bii**). An initial increase in delta power shortly after stimulus onset, due to the environment, can be observed, followed by decreases in the frequency range of theta and beta. Between 2000 and 800ms before the decision to leave, there is a decrease in delta power (**Ci, Cii**). Gray areas and black contours show significant clusters after permutation-based correction for multiple comparisons.

3.3.2.3 Beta ramps

Because we saw a decrease in beta activity the better the patch was from both the stimulus-onset perspective and the response-locked perspective, we averaged the beta frequencies of the raw time-frequency data and plotted them in patch clusters (***Figure 3-5 Ai; Aii***). Effects were most pronounced in the parietal areas and contralateral to motion, so here we show electrodes Pz and C3. From 500ms after stimulus onset, high-

reward patches show a stronger decrease in beta power compared to low-reward patches, whereas this effect seems to change direction as the decision to leave approaches, with a stronger decrease for low-reward patches. However, we cannot prove that this stimulus-onset-locked beta represents the same phenomenon as the response-locked beta. Furthermore, we split the group in participants with good and poor efficiency based on their deviation from optimum and plot the averaged beta frequencies for both groups (***Figure 3-5 Bi; Bii***). Participants with poor efficiency seem to have a stronger decrease in the beta frequency range both, after stimulus onset and close to the decision to leave. These results are rather exploratory in nature and serve to better describe the effect found in the beta frequency range, as we did not have strong prior hypotheses about an association between midfrontal beta and foraging.

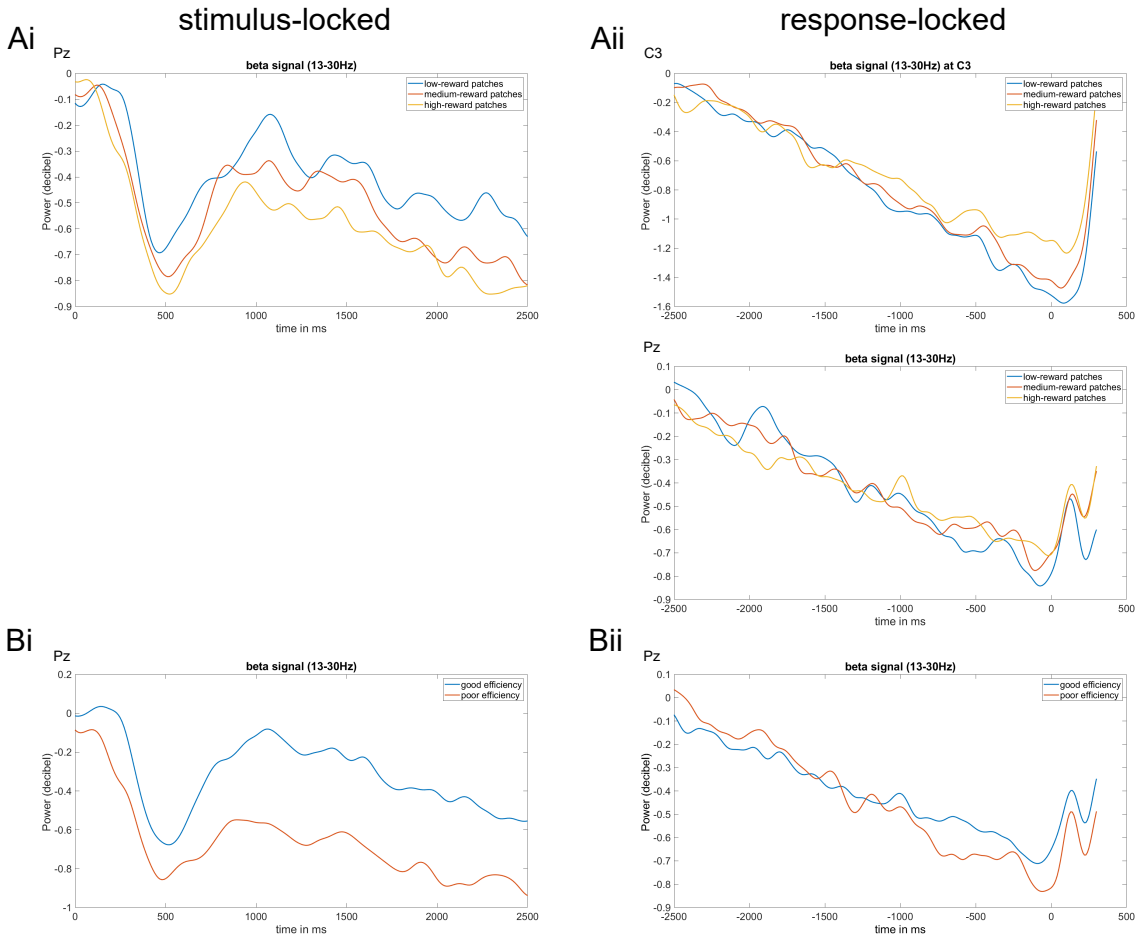

Figure 3-5. Raw beta signaling due to quality of patch and efficiency of participants. The left part of the figure shows the stimulus-onset locked results, while the right part shows the response-locked results. Beta power appears to be lower for patches yielding less reward compared to higher quality patches starting 500ms after the onset of harvesting (**Ai**). This effect

(Figure 3-5 continued) seems to reverse just before the decision to leave is executed (**Aii**). Beta power appears to be lower in participants showing poor efficiency compared to high-performing participants, especially in the initial harvest phase (**Bi, Bii**).

3.4 Discussion

The aim of the present study was to examine the temporal neurophysiological dynamics underlying decision-making during foraging in humans. Consistent with expectations, the behavioral results demonstrated that participants' behavior approximated that of an optimal agent, as proposed by the MVT. However, participants' performance deviated from the optimal behavior by a certain extent, and they displayed systematic biases when adapting to specific manipulations. We observed suboptimal adaptation to a resource-poor environment compared to a resource-rich one, consistent with previous evidence (Garrett & Daw, 2020). In contrast to prior findings (Constantino & Daw, 2015; Harhen & Bornstein, 2023), participants in this study generally "underharvested" the available options. This discrepancy may be attributed to the nature of the paradigm, which induced a sense of urgency with time-related information, unlike other paradigms that operate differently. In alternative patch-leaving paradigms (Constantino & Daw, 2015; Harhen & Bornstein, 2023; Wittmann et al., 2016), the decision to stay/leave did not disrupt the continuous harvest, as seen in our study, where the decision timing *did* influence the harvesting duration itself. Instead, the stay/leave choice occurred sequentially after the harvest phase was completed. While participants exhibited stronger sensitivity to changes in the foreground reward rate (FRR) and adapted their choices on a trial-by-trial basis accordingly, they did not adequately consider changes in the background reward rate (BRR) associated with changes of the environment. Notably, the adaptations to changes in the FRR were insufficient in magnitude. Participants viewed the accumulated reward on the screen and had to compute the derivative of this value to determine the optimal time to stop foraging (iRR). This complex task could have contributed to suboptimal behavior. Alternatively, it is plausible that the optimal adaptive approach proposed by the MVT is too complex and inefficient for humans. Instead, humans may rely on simpler heuristics to make decisions that align with a basic cost-benefit calculation (Findling et al., 2021).

On the physiological level, we were able to identify distinct signals for different aspects of the foraging decision. The BRR operationalized as a change of environment has a sustained impact, because it does not change on a trial-by-trial basis. The signal tracking the BRR is represented in form of a more positive EEG amplitude when the

environment is poor compared to rich and an increase in the frequency range of delta when the harvest phase begins, which turns into a decrease approximately 2 to 1 sec before the decision to leave is executed. Cavanagh (2015) has already proposed a mechanistic role for delta band activities in motivating action selection. Moreover, synchronous activity in the delta frequency band was associated with the coordination of distant cortical networks during decision-making (Nácher et al., 2013). If that is true for the present delta activities, this result could indicate that a network that goes beyond the aMCC is involved in foraging decisions. This would also explain why, contrary to assumptions, we found a strong association with parietal activities in addition to frontal activations.

After one second post stimulus onset, the delta increase is alternated by a decrease in the theta frequency range. A decrease in theta seems to represent a shared component, which plays a role for tracking the BRR as well as the FRR. Here, we see a decrease in the theta frequency range 1500ms after stimulus onset and again early, about 2.5 sec before, until the decision to leave is executed. This reduction in midfrontal to parietal theta activity is associated with both the environment and the trial or patch-related scale. This finding is consistent with Jacobs et al. (2006), who found a relationship between midfrontal theta and decision-making, showing a decrease in preresponse theta power. We expected to find theta band oscillations involved in foraging decisions, since previous evidence associated it with strategic cognitive control in order to reduce uncertainty (Cavanagh et al., 2012; Cavanagh & Frank, 2014; Cavanagh & Shackman, 2015).

Consistent with our expectations, we observe broad decreases in beta activity over parietal and contralateral regions, mainly as a function of the FRR, starting from 500ms after the onset of the harvest phase and persisting until the decision to leave is made. Notably, the higher the patch quality becomes, the less beta activity occurs, and the later the participants leave the patch. Usually, decreases in beta power are related to motor preparation (Khanna & Carmena, 2015). However, in this context, the decreasing beta activity linked to improved patch quality appears to be intricately tied to decision processes rather than being solely a motor-related phenomenon. This leads us to the conclusion that the reduction in beta power does not merely reflect motor preparation but indeed encapsulates a decision variable (Fischer et al., 2018; Rogge et al., 2022). Specifically, within the patch-leaving framework, the variable corresponds to

patch quality, directly determining the timing of leaving. This conclusion gains further support when comparing results with a task similar to the present paradigm, but without a foraging requirement – purely motor. In this task, a pre-response reduction in beta power is also evident, but it is not correlated with patch quality (***Figure 3-6 Di; Dii***).

Beta and theta signaling have previously been shown to be associated with the accumulation of information until a response threshold is crossed (Donner et al., 2009; Fischer et al., 2018; Guan et al., 2023; O'Connell & Kelly, 2021; Pape & Siegel, 2016; Rogge et al., 2022). Moreover, MVT predicts the ideal time to leave a particular option when the FRR falls below the BRR, thus defining a threshold. The FRR-associated reduction in beta and theta activity suggests that these neuronal signals serve as inhibitory signals until a certain threshold is reached. This interpretation is supported by the finding that beta signals are most negative at the onset of harvest for high-reward patches, whereas a reversal occurs just before the decision to leave, with the beta signal most negative for low-reward patches, thus possibly marking a threshold. It remains unclear whether the slope leading up to the threshold becomes less steep when patch quality is better compared to worse *or* whether the threshold itself increases as a function of a peak in beta or theta reduction.

Enhanced patches necessitate longer periods of exploitation. As patch quality improves, participants indeed tend to extend their stay, but this adaptive response is not proportionate, resulting in more suboptimal behavior as patch quality increases. This phenomenon of less-optimal behavior in correlation with higher patch quality might find its explanation in the context of collapsing decision thresholds as time elapses, as recently proposed by Kirschner et al. (2023). Furthermore, we found evidence of a relationship between the beta signal and participants' efficiency, i.e., how close they approach the optimum of a model. Poor efficiency could be related to a smaller decrease in the beta signal, which would imply lower inhibition or a lower threshold, resulting in an earlier departure than the optimum suggested by the model. The extent and mechanisms through which the reduction in beta and theta power impacts foraging decision-making by affecting decision thresholds warrant further investigation in subsequent studies. Replication of the findings presented in this work is essential to validate these relationships.

As mentioned above, to better distinguish between neuronal phenomena related to purely motor processes and more cognitive processes inherent in foraging decision-making, we used a very similar task to our paradigm. This control task was designed to emphasize more execution rather than foraging demands. As illustrated in the supplementary material (***Figure 3-6***), in the time-frequency decomposition of the raw data reveals notable activations at the onset of the harvest phase and before the keystroke occurs. Increased activity in the frequency range of delta and theta emerges at the beginning of the harvest phase, which could be interpreted within the context of a P300. However, the signal pattern of this motor-related task differs significantly from that observed in the foraging paradigm, particularly concerning the reflection of decision-related variables. Remarkably, there is no longer a reduction in theta or beta power associated with patch quality. In essence, this comparison supports the interpretation that the observed phenomena are neural processes closely tied to foraging decision-making, rather than being predominantly influenced by motor-associated factors.

In conclusion, the aim of the present study was to investigate the temporal neurophysiological dynamics of foraging decisions in humans. Our findings are the first to describe distinct patterns of beta, theta, and delta oscillations that relate to foraging behavior. Delta band activities appear as a unique signal of BRR with a maintained impact. Theta band activities represent a shared component, reflecting both FRR and BRR. Theta and beta band activities seem to continuously track the foraging process, beginning as early as the initial harvest phase leading up to the decision to leave a particular patch. Future studies should focus on continuous mapping of the foraging process, examining the nature of any threshold or slope adaptation. The use of electrophysiological signals is an eminently suitable approach to better describe foraging decisions, and we encourage further research to contribute to this still highly neglected opportunity to improve the understanding of neuronal mechanisms of complex human decision-making.

3.5 Supplemental Information

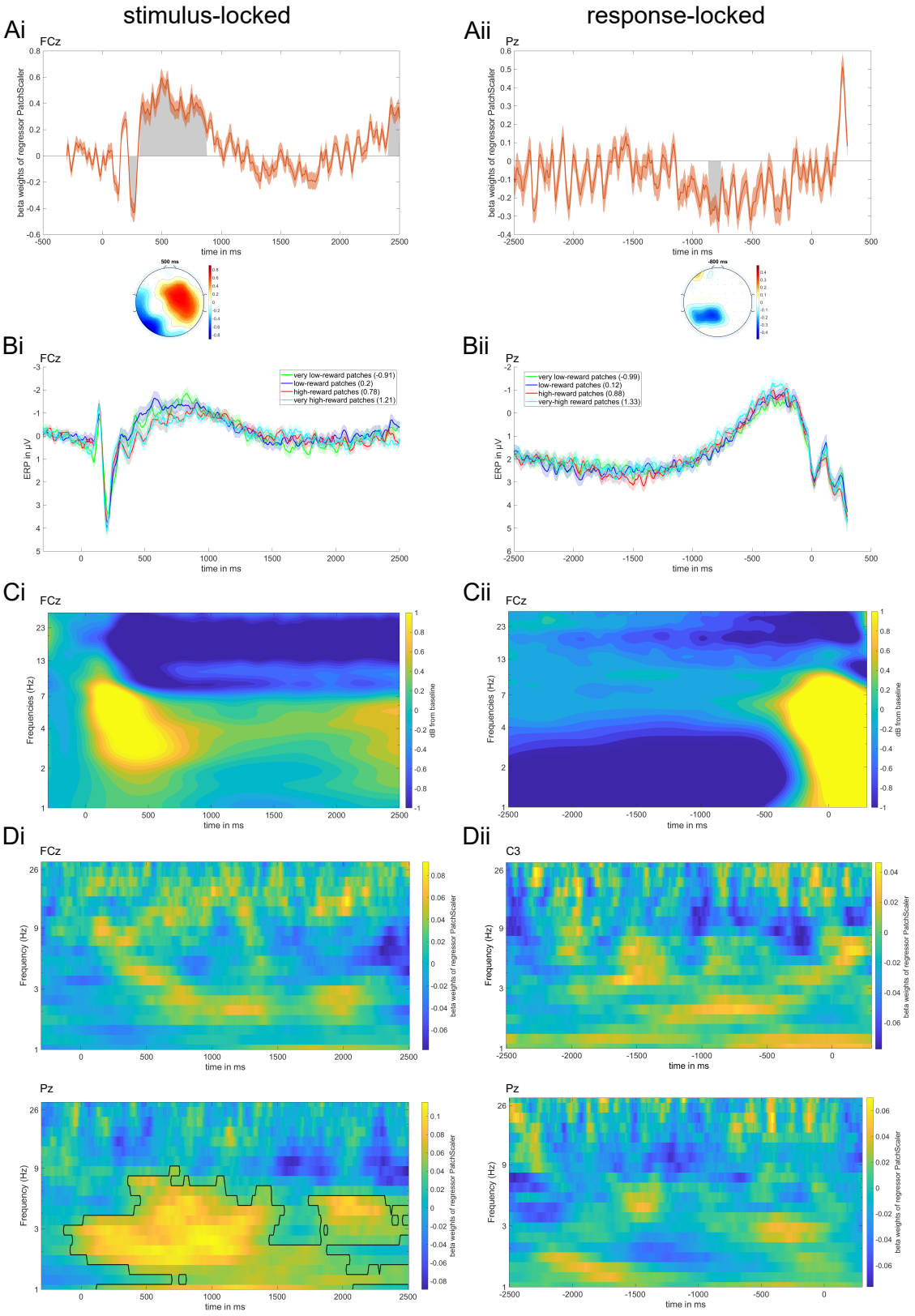

Figure 3-6. Effect of PatchScaler onto the EEG signal for the motor task and raw time-frequency decomposed data. The left part of the figure shows the stimulus-onset

(Figure 3-6 continued) locked results, while the right part shows the response-locked results of the regressor PatchScaler. The signal is more positive, the better the patch quality mostly in the beginning of the harvest phase (**Ai, Aii, Bi, Bii**). Time-frequency decomposed data shows an initial increase in theta power followed by a sustained decrease in beta power (**Ci**). Before the decision to leave is executed, a sustained decrease in delta power can be observed, before this effect reverses just before the decision is made (**Cii**). Regression in time-frequency space reveals a broad increase in lower frequencies in the initial harvest phase (**Di**), while a decrease in delta/theta power before the decision to leave is observed (**Dii**). Gray areas and black contours show significant clusters after permutation-based correction for multiple comparisons.

The following chapter is based on an article published in NeuroImage:

Kirsch, F.[1], Kirschner, H.[1], Fischer, A. G., Klein, T. A.[2], & Ullsperger, M.[2] (2022). Disentangling performance-monitoring signals encoded in feedback-related EEG dynamics. *NeuroImage*, *257*, 119322. https://doi.org/10.1016/j.neuroimage.2022.119322

1 These authors contributed equally to this work.

2 TK and MU should be considered joint senior author.

4 Study II: Disentangling performance-monitoring signals encoded in feedback-related EEG dynamics

4.1 Introduction

In general, feedback is important for learning and adaptive, goal-directed behavior. When feedback informs about an action outcome, a feedback-locked sequence of EEG-components consisting of a frontocentrally distributed feedback-related negativity (FRN; Miltner et al., 1997), the frontocentral P3a, and a parietal P3b can be observed (Ullsperger, Fischer, et al., 2014). The reinforcement learning (RL) theory of Holroyd and Coles (2002) states that the amplitude of the FRN correlates with the reward prediction error (RPE). Indeed, it is proposed by several researchers that the FRN encodes an RPE (Chase et al., 2011; Cohen & Ranganath, 2007; Holroyd & Coles, 2002; Holroyd & Krigolson, 2007; Nieuwenhuis et al., 2004; Sambrook & Goslin, 2015; Ullsperger, Danielmeier, & Jocham, 2014), having a stronger deflection when outcome expectation is violated (Chase et al., 2011; Holroyd & Krigolson, 2007; Walsh & Anderson, 2012). More negative RPEs are suggested to be associated with a stronger posterior mesial frontal cortex (pMFC) response (Jocham et al., 2009) and a larger FRN. According to the RL-theory, unexpected negative outcomes should elicit larger FRNs than expected negative outcomes (San Martín, 2012). Numerous studies have indicated that the FRN amplitude scales with a "signed" RPE: In the case of a worse-than-expected outcome (negative RPE), a strong FRN is elicited, whereas a smaller and weaker FRN is observed after better-than-expected outcomes (positive RPE; Fischer & Ullsperger, 2013; Hajcak et al., 2007; Holroyd & Coles, 2002; Walsh & Anderson, 2012). Expectations are generated by experience and incorporate global values (frequent/infrequent) but can dynamically adjust to recent events, such as local surprise

generated by trial micro-structures (Holroyd & Coles, 2002).

Beyond the RL-account, there are alternative approaches suggesting the FRN corresponds to an unsigned RPE signal which is sensitive to unlikely and therefore salient events indicating surprise (independent of the direction of the expectedness violation, Alexander & Brown, 2011; Donkers & van Boxtel, 2005; Ferdinand et al., 2012; Hauser et al., 2014; Talmi et al., 2013; Walentowska et al., 2019; Yeung et al., 2005). In contrast, some studies show a dependence of the FRN amplitude on valence and postulate that the magnitude of probability is represented later in more parietal components like the P3b (Kamarajan et al., 2009; Sato et al., 2005; Toyomaki & Murohashi, 2005; Yeung & Sanfey, 2004). Since the P3 complex is increased upon low probability events (Johnson, 1986), FRN and P3 correlate with the common factor surprise and therefore, it is possible that the FRN could be overlapped by the P3 (Walsh & Anderson, 2012). Recently, it has been proposed that the FRN effect of being larger for negative as compared to positive feedback is actually driven by a positive deflection following positive outcomes (RewP; Baker & Holroyd, 2011; Foti & Weinberg, 2018; Holroyd et al., 2008; Krigolson, 2018; Proudfit, 2015). Unexpected outcomes are supposed to elicit the ERP component N200, while trials with unexpected rewards elicit a feedback-related positivity (RewP) and, in consequence, the RewP overshadows the effect of the N200. However, given that the positive (RewP) and negative (FRN) deflections overlap in time, it remains unclear which of them captures systematic changes in reward processing best (Gheza et al., 2018). Some authors, in fact, seem to suggest that FRN and RewP represent the same EEG phenomenon, just with opposite sign (Krigolson, 2018; Proudfit, 2015). In this context, we would like to emphasize that the interpretation of results depends on the definition and quantification of the FRN. While some authors quantify the FRN as the loss-minus-win difference wave or the N2-like component following feedback, others suggest a difference between feedback condition-specific components to loss and gain (Cavanagh et al., 2019). In the present study, we are interested in factors that independently influence feedback-related EEG dynamics in the latency range of the FRN (and P300). Therefore, we applied a single-trial regression approach instead of using an ERP quantification method to avoid interpretation issues arising from different approaches. We merely refer to the FRN and P3a/b to guide the reader in terms of latency and topography during which variables of interest modulate the stereotypical

ERP sequence after visual feedback (Ullsperger, Fischer, et al., 2014).
Concerning learning from feedback, larger behavioral adjustments were found after participants received negative compared to positive feedback, which was also reflected in the amplitude of the FRN (Holroyd & Krigolson, 2007). In the same study, the interaction of expectedness and valence of feedback was associated with the extent of behavioral adaptation.

Whereas the FRN is suggested to reflect an early evaluation process involving the calculation of a prediction error, the P3 has been proposed to translate this information into attentional and working memory processes, and to initiate behavioral adaptation (Donchin & Coles, 1998; Polich, 2007; Verleger, 1997; Verleger et al., 1994). The P3 complex consists of two positive ERP deflections, P3a and P3b, which are elicited by potentially action-related stimuli (Ullsperger, Fischer, et al., 2014). The early frontocentral P3a seems to reflect fast orienting and stimulus-driven attention mechanisms (Kirschner et al., 2022; Ullsperger, Fischer, et al., 2014), whereas the more sustained parietal P3b has been proposed to be associated with surprise (Donchin & Coles, 1998; Mars et al., 2008) and action value updating (Ullsperger, 2017; Ullsperger, Fischer, et al., 2014). It is typically found that unexpectedness or negative valence of the feedback give rise to a larger P3b than expected or positive feedback (de Bruijn et al., 2004; Fischer & Ullsperger, 2013; Walentowska et al., 2016). Nevertheless, some sources report no valence effects (Yeung & Sanfey, 2004) or even reverse findings concerning valence (Hajcak et al., 2007; Severo et al., 2018).

Since there is inconsistent evidence about the influence of valence and expectedness (San Martín, 2012) and their interaction on the neural response to outcome processing, we approached this question differently from former studies, many of which used difference waves (Glazer & Nusslock, 2021; Hajcak et al., 2007; Holroyd & Coles, 2002; Holroyd & Krigolson, 2007; Holroyd et al., 2008; Talmi et al., 2013; van Boxtel, 2004; Walsh & Anderson, 2012). In the present study, we aimed at giving a holistic perspective on feedback processing. As noted above, previous work has hinted that both surprise and valence contribute to feedback related EEG dynamics, but the precise nature of their interaction remains elusive. Here, we leveraged single trial regression and the power of a large sample to parse out the contributions of these

factors on the EEG signal. While we have previously used instrumental learning tasks (Burnside et al., 2019; Fischer & Ullsperger, 2013; Kirschner et al., 2022), we exclude this learning aspect here. Moreover, when we use the term "FRN" in the context of the present work, we mean modulations of the EEG signal in a latency range of 200 to 300ms over frontocentral regions, and do not refer to any particular ERP quantification. Our goal was to differentiate changes of the FRN amplitude and the P3 complex as a function of valence and expectedness within a large sample. We therefore systematically manipulated valence and expectedness of the outcome to investigate the influence on the neuronal signal with the help of a single-trial regression approach. Feedback valence was either positive or negative. Outcome expectedness was conceptualized as global surprise. Here, the level of task difficulty was manipulated between blocks by increasing the expectancy of positive feedback or increasing the expectancy of negative feedback without the subjects' knowledge. Additionally, we investigated local surprise by examining the influence of the recent trial history on the feedback-locked neuronal signal. If the FRN reflects mere surprise, a response to salience in the form of an unsigned RPE, the component should be insensitive to valence. Therefore, a component within the timeframe of the FRN should show no main effect of valence and no interaction of expectedness and valence, but a strong main effect of unsigned RPE size involving outcome probability (Sambrook & Goslin, 2015). In consequence, the FRN should be equally large for worse-than-expected and better-than-expected outcomes (Walentowska et al., 2019). Contrarily, if the FRN encodes signed RPEs, we would expect a valence x expectedness interaction, where the FRN for unexpected events differs between worse-than-expected and better-than-expected events. A sole main effect of valence would mean that the FRN represents the outcome itself and does not encode any RPE. Two non-interacting main effects of valence and expectedness would indicate two independent processes influencing the neuronal signal in the FRN latency range. Hypotheses are visualized in **Figure 4-1 C**. Concerning the P3 complex, we expect a clear effect of expectedness: unexpected outcomes should induce a larger P3b than expected outcomes.

To elaborate on the topic of behavioral adaptations following feedback, we used an advanced measure of change in response time between two consecutive trials, which reflects improvement or deterioration in task performance from one trial to the next.

Since previous research has found that the type and expectedness of feedback could affect the extent of adaptation and that this adaptation has correlates in the neuronal signal (Holroyd & Krigolson, 2007), we considered the influence of the change in response time on the neuronal signal as a function of the valence and expectedness of the feedback. Specifically, we expected larger adaptations after negative feedback. Exploratorily, this adaptation could be affected by outcome expectancies, whereby unexpected negative feedback should be accompanied by larger and expected negative feedback by smaller behavioral adjustments.

4.2 Methods

4.2.1 Participants

1000 young, healthy participants were recruited at the Radboud University of Nijmegen, Netherlands (388 datasets; 03/2011-06/2011) and at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, Germany (all remaining; 05/2012-03/2016). Screened via interview, exclusion criteria were: any present or past psychiatric or neurological disorders, regular use of medication, drug abuse, alcohol intake at day of study. 8 subjects had to be excluded due to recording failures or poor data quality (see 2.3 for details). The sample consists of 493 female and 499 male subjects between 18 - 40 years (*M* = 24.22; *SD* = 4.03). The majority of the participants were right-handed (N = 914), 44 were left-handed, and 33 ambidextrous or retrained (1 not reported). The study was approved by the Institutional Review Board/Ethics Committee of the Radboud University of Nijmegen (ECG04032011) and the University of Leipzig (285-09-141209) and written informed consent was obtained from all participants after briefing prior to study enrolment. The study was conducted in accordance with the Declaration of Helsinki.

4.2.2 Experimental paradigm

We used a modified version of a time estimation task, where the participants had to estimate the duration of one second by keystroke with a fixation cross as onset time point (Gruendler et al., 2011; Holroyd & Krigolson, 2007, **Figure 4-1 Ai**). Positive Feedback for a correct response (i.e. responding in a time window of 1000ms ± 100ms initially) was given in form of a green smiley, negative feedback for an incorrect response was visualized with a red frowny. Within the experiment, every subject underwent three conditions, where the time window (TW) for correct responses was adapted differently such that negative feedback was more likely than positive feedback ("difficult"), negative and positive feedback were equally likely ("control"), and negative feedback was less likely than positive feedback ("easy"), respectively. Within the control condition, the TW for positive feedback increased by 10ms on error trials and decreased by 10ms on correct trials symmetrically. In the easy condition, the window size increased by 12ms on error trials and decreased by 4ms on correct trials. In the difficult condition, adaptation reversed compared to the easy condition, so that the TW narrowed faster

on correct responses (-12ms) than it grew after incorrect responses (+4ms; **Figure 4-1 Aii**). Participants always started with the control condition, where a TW for positive feedback was initialized at 1000ms ± 100ms (positive feedback was given when the response fell between 900 and 1100ms). They then continued with either the easy or the difficult condition and went through each condition once, counterbalanced over participants. The time window to start with in the first trial of a new block was equal to the time window resulting from the last trial of the preceding block. The control condition only served to intercept initial adaptation processes and to define the initial time window and was therefore not equivalent to the other conditions, which is why it was excluded from analysis. The task consists of 450 trials in total, 50 trials in the control condition, 200 trials each in the difficult and easy conditions. The inter-trial interval (feedback-cue interval, FCI) varied between 750ms and 1250ms. **Figure 4-1 B** shows the conceptual mapping of the task blocks (easy/difficult) to the associated size of the RPE and the valence of the outcome.

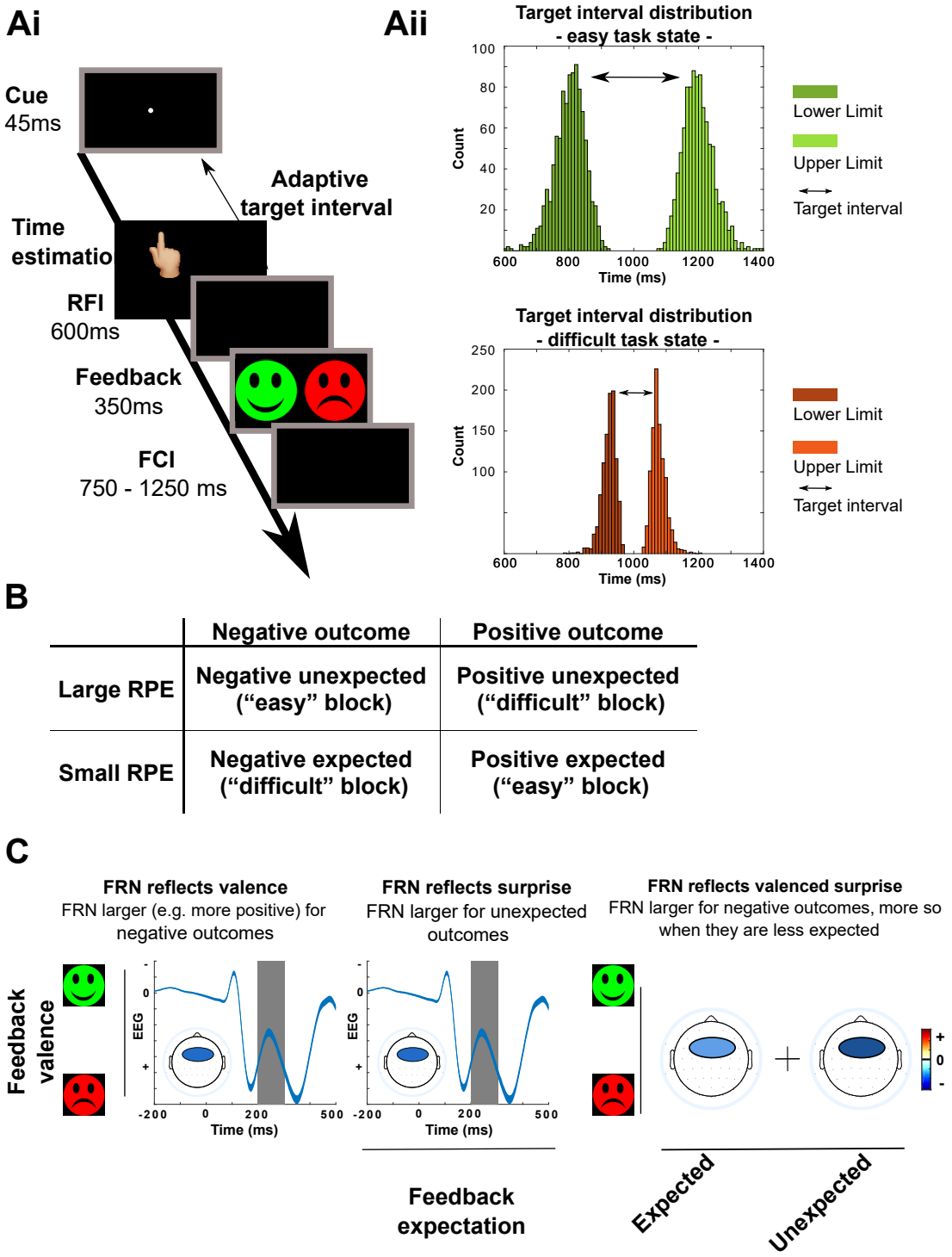

Figure 4-1. Illustration of the task design and theoretical hypotheses on task factors on the FRN signal. (**Ai**) Timeline of a single trial. Each trial starts with a cue represented by a fixation cross. From cue onset on, participants are instructed to estimate one second and to indicate the end of the estimated time interval by button press. The time window for a correct response is adaptive and differs between conditions (control, easy, difficult). A reaction feedback interval (RFI) of 600ms follows the response. Participants then receive either positive or negative

(Figure 4-1 continued) feedback in the form of a green smiley for correct response and a red frowny for incorrect response, which is displayed for 350ms. This is followed by a variable feedback cue interval (FCI) of 750-1250ms before the next trial starts. (**Aii**) Target interval distributions for the different task states (easy and difficult). (**B**) Experimental conditions and their respective correspondences to a reward prediction error (RPE) as a function of the valence of the feedback. (**C**) Illustration of theoretical hypotheses for FRN signals. Left: FRN reflects valence. FRN is larger (e.g., more negative) on all negative outcomes. When comparing positive vs. negative feedback, this should be reflected in frontocentral negativity as depicted in the respective topography plot. Middle: If the FRN is reflecting mere surprise, the FRN should be larger for unexpected feedback (positive outcomes in the difficult task state and negative outcomes in the easy task state). Right: The FRN reflects "valenced surprise" (e.g., prediction errors). Larger FRN in negative outcomes (reflected in a frontocentral negativity), more so when they are less expected than when they are common.

4.2.3 EEG acquisition and processing

Electroencephalic signals were continuously recorded at 500Hz sampling rate with BrainAmp MR plus amplifiers (Brain Products) from 64 Ag/AgCl sintered electrodes, which were mounted in an elastic cap according to the extended 10-20 system with impedances kept below 5kΩ. The ground electrode was placed at the sternum. Electrodes to capture horizontal and vertical eye movements were mounted next to both eyes and above and below the left eye. The signal was online referenced to A1 (left mastoid). The recorded data was high (0.5Hz) - and low (30Hz) -pass filtered, re-referenced to common average, and epoched from - 400ms to 1500ms locked to feedback onset. Artifactual epochs were automatically rejected based on signal outliers. Epochs that deviate over 5 SD from the mean probability distribution of the EEG signal were excluded (Delorme et al., 2007). We specified that a minimum of 10 trials but no more than 10% of the trials (N = max. 45) should be rejected. Therefore, the initial threshold of 5 SD was adaptively increased or decreased with a step size of 0.1 SD. This resulted in an average rejection of 20 epochs across all participants (range N = 10-44). Epochs were then demeaned and submitted to adaptive mixture independent component analysis (AMICA, Palmer et al., 2012). Independent components including artifactual signals (i.e. eye blinks) were rejected with the help of sample-based ratings of two EEG-experienced researchers in combination with a correlation-based approach (inspired by the Corrmap approach; Viola et al., 2009). A baseline of − 350ms until 0ms prior to feedback onset was used. The data was then analyzed with multiple robust single-trial regression analyses (see Fischer & Ullsperger, 2013 for details). EEG datasets for which ICA did not converge or too less trials for specific regressors existed were

excluded (8 subjects), resulting in a final sample of 992 participants. For EEG and behavioral analysis, EEGLab 13.5 toolbox (Delorme & Makeig, 2004) and customized code written in MATLAB R2019b version 9.7.0.1190202 (MathWorks) was used.

4.2.4 Data analysis

4.2.4.1 Behavioral analyses

Data for building behavioral regressors were either directly derived from the behavioral data or, in some cases, calculated from combinations of other behavioral variables. *Valence* (Val) represents the qualitative direction of the feedback, either positive (=0) or negative (=1). The regressor *expectedness* (Exp) is a dichotomous variable, where unexpected events (=0) indicate trials with negative feedback during the easy condition and trials with positive feedback during the difficult condition. Expected events (=1) are coded vice versa (positive feedback in easy condition; negative feedback in difficult condition). *Reaction time* (RT) reflects the absolute reaction time from fixation cross onset until button press. The following variables reflect local surprise: the *number of trials since the last negative or positive event* (TrialsSinceNeg/Pos) indicates how long ago the last event of the same valence occurred. The longer it has been, the bigger the local surprise. For these two variables, corresponding trials are cumulated separately for negative and positive events. In addition, if participants build up an internal representation of the target time, another form of local surprise can be conceptualized as the absolute difference between 1000ms (the target time) and trial-based reaction time. This variable is represented by the parametric regressor *reaction time deviation* (RT_dev; unsigned). The closer the participant is to 1000ms and still receives negative feedback, the bigger the local surprise. *Reaction time change* (RT_change) was calculated (Equation 1) by subtracting the absolute reaction time difference of the following trial from the absolute reaction time difference of the current trial and therefore, represents a measure of performance adaptation. The results of this calculation (see Equation *4-1*) could either take on positive values, which represent performance improvement (i.e. coming closer to 1000ms), or negative values, which represent performance deterioration (increasing distance to 1000ms):

$$RT\_change = |(RT_{current\ trial} - 1000ms)| - |(RT_{following\ trial} - 1000ms)| \quad (4\text{-}1)$$

As noise regressors without interest, logarithm of the trial number in the current block (BlockTr) and inter-trial interval (FCI) were included in all regression models. The parameters BlockTr, FCI, RT, RT_dev, TrialsSinceNeg/Pos, and RT_change were z-standardised.

We calculated a robust multiple regression with behavioral adjustment (absolute difference between two consecutive trials in ms) as the outcome. Valence, expectedness, their interaction, as well as BlockTr and FCI served as predictors:

$$Behavioral\ adjustment = \beta_0 + \beta_1 \times Val + \beta_2 \times Exp + \beta_3 \times BlockTr + \beta_4 \times FCI + \beta_5 \times Val \times Exp + \varepsilon \quad (4\text{-}2)$$

4.2.4.2 Grand averages

We calculated grand average ERPs at two electrodes for crossed conditions (negative feedback-expected; negative feedback-unexpected; positive feedback-expected; positive feedback-unexpected) by averaging across all subjects. The site of maximal FRN activity is, according to the literature (Williams et al., 2021), electrode FCz and the site of maximal P3b activity is, according to the literature (Intriligator & Polich, 1994; Polich, 2007), electrode Pz. Therefore, we chose FCz and Pz to visualize the neuronal signal depending on feedback.

4.2.4.3 Single-trial EEG analyses

We furthermore employed several multiple robust regressions, within (1st level) and across subjects (2nd level; (Fischer et al., 2016; Fischer & Ullsperger, 2013). General linear models (GLM) were built to regress single-trial EEG activity at each electrode and time point against behavioral parameters. The regressions were performed on 59 electrodes in a time window from -200ms to 1000ms, feedback-locked. The output of these analyses was in the form of regression coefficients revealing the time course and scalp topographies of the relationship between each predictor and neuronal activity. Standardized beta-values can be tested via two-tailed one-sample t-tests, which were done separately at each data point in a whole-brain approach across subjects. To account for multiple comparisons, p-values within one model were corrected using false discovery rate (FDR). Trials of the control condition were excluded from regression analyses.

Within the first GLM (1a), we were interested in the influences of valence of feedback, expectedness, and the interaction of both on the neuronal signal:

GLM 1a: $EEG\ amplitude = \beta_0 + \beta_1 \times Val + \beta_2 \times Exp + \beta_3 \times BlockTr + \beta_4 \times FCI + \beta_5 \times Val \times Exp + \varepsilon$ (4-3)

To further disentangle the results revealed in the first GLM, we split the data in expected and unexpected trials within a subordinate GLM (1b). This enables us to have a more detailed look on the effect of valence comparing negative- and positive-feedback-trials for expected and unexpected trials separately:

GLM 1b: $EEG\ amplitude = \beta_0 + \beta_1 \times Val + \beta_2 \times BlockTr + \beta_3 \times FCI + \varepsilon$ (4-4)
(ran separately for expected and unexpected outcomes)

By building a second main GLM and splitting the data in positive- and negative-feedback trials, it was possible to resolve the interaction between valence and expectedness and to further investigate the differential effects of expectedness or global surprise, respectively. Additionally, we included further variables, which give information on local surprise, like TrialsSinceNeg/Pos and RT_dev:

GLM 2: $EEG\ amplitude = \beta_0 + \beta_1 \times Exp + \beta_2 \times RT + \beta_3 \times TrialsSinceNeg/Pos + \beta_4 \times RT\_dev + \beta_5 \times BlockTr + \beta_6 \times FCI + \varepsilon$ (4-5)
(ran separately for trials with negative and positive feedback)

In the last GLM 3, we were interested in the neuronal signal changes due to performance adaptation. Therefore, we created main GLM 3a including the regressor RT_change, which reflects performance improvement or deterioration between the current and the consecutive trial. The interaction of feedback valence and RT_change served as a predictor as well (Eq. *4-6*). In a subordinate GLM 3b, we disentangled this interaction by running separate regressions for trials with negative and positive feedback (Eq. *4-7*). Because the participants tend to adjust their RT more after negative feedback, we calculated another subordinate GLM 3c for negative-feedback trials to examine the interaction between expectedness and behavioral adaptation exploratorily (Eq. *4-8*).

GLM 3a: $EEG\ amplitude = \beta_0 + \beta_1 \times Val + \beta_2 \times RT\_change + \beta_3 \times Exp + \beta_4 \times BlockTr + \beta_5 \times FCI + \beta_6 \times Val \times RT\_change + \varepsilon$ (4-6)

GLM 3b: $EEG\ amplitude = \beta_0 + \beta_1 \times RT\_change + \beta_2 \times Exp + \beta_3 \times BlockTr + \beta_4 \times FCI + \varepsilon$ (4-7)
(ran separately for trials with negative and positive feedback)

GLM 3c: $EEG\ amplitude = \beta_0 + \beta_1 \times RT\_change + \beta_2 \times Exp + \beta_3 \times BlockTr + \beta_4 \times FCI + \beta_5 \times Exp \times RT\_change + \varepsilon$ (4-8)
(ran for trials with negative feedback)

We were particularly interested in beta-value peaks approximately in the latency ranges (FRN: 200-300ms; P3a/P3b: 300-600ms; P3a usually earlier) and topographic locations (FRN/P3a: frontocentral; P3b: centroparietal) of FRN, P3a and P3b as reported in the literature (Polich, 2007; Sambrook & Goslin, 2015; San Martín, 2012; Ullsperger, Fischer, et al., 2014), because these are components that typically occur following feedback. Thus, for post-hoc t-tests after regression analysis, we derived latencies of local beta-value peaks of the predictors in FRN- or P3-time windows based on visual inspection and comparing with the respective ERP. Individual beta-value peak time points (not all shown in ***Figure 4-4***) were extended by including beta-values in a time window of ±20ms around the peak and using the average of them. The indicated electrodes usually include the electrode with the strongest effect, otherwise they represent strong local effects.

4.2.5 Data and code availability

The conditions of our informed consent form do not permit public archiving of the raw data because participants did not provide sufficient consent. Researchers who wish to access processed and anonymized data from a reasonable perspective should contact the corresponding author. Data will be released to researchers if it is possible under the terms of the GDPR (General Data Protection Regulation). The code of the toolbox we used for the regression analysis can be found here: http://www.adrianfischer.de/teaching.html.

4.3 Results

4.3.1 Behavioral results

Overall, the participants performed well in the task, the mean reaction time across all conditions was 1032.10ms ± 218.21ms (***Figure 4-2* A**). Manipulation of difficulty by different task states was successful: within the easy condition, participants received positive feedback in 70.27% (SD_{easy} = 3.98%) of the trials, while in the difficult condition, only 29.62% ($SD_{difficult}$ = 2.50%; $M_{control}$ = 41.70%; $SD_{control}$ = 8.18%) of the trials were followed by positive feedback (***Figure 4-2* B**).

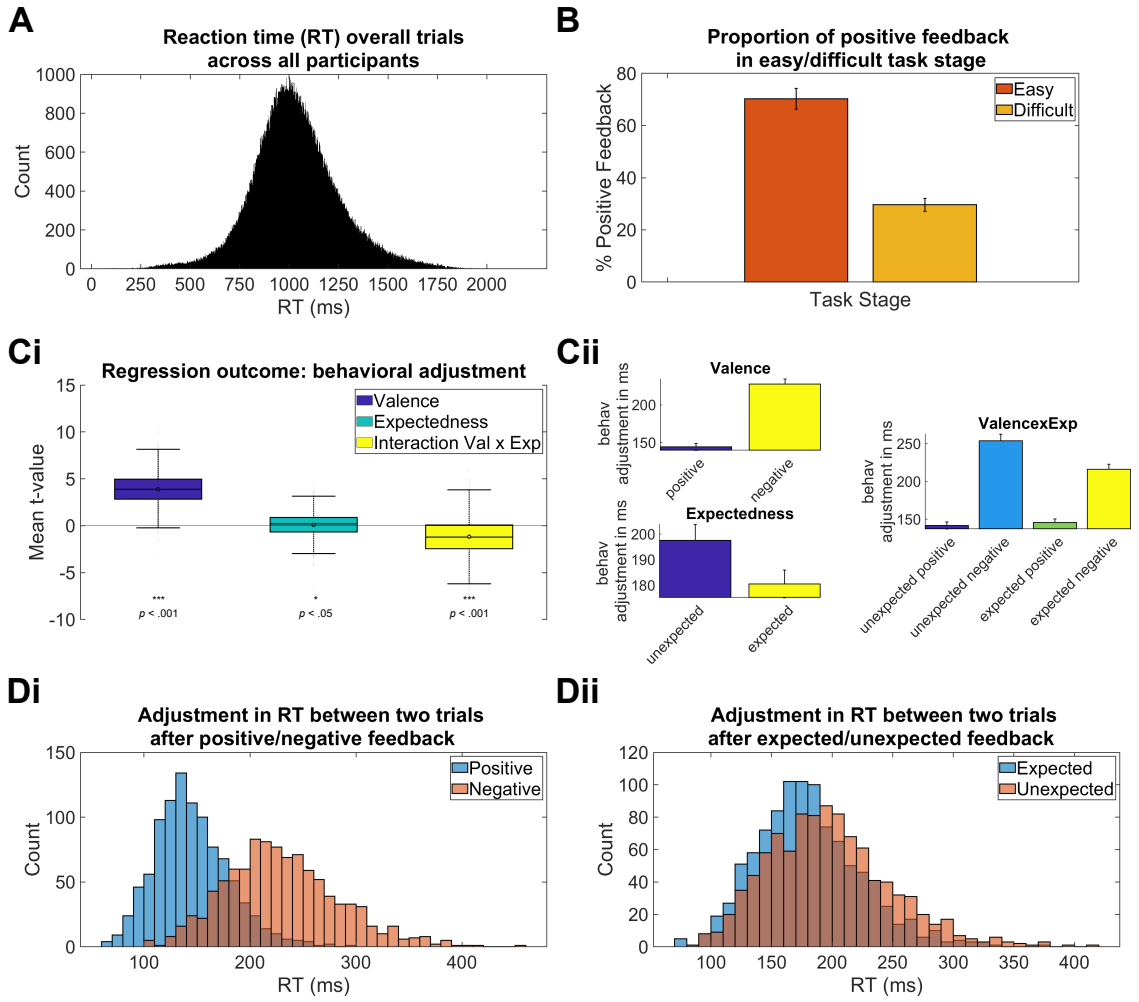

Figure 4-2. Frequencies of reaction times and behavioral adjustments and influencing factors. (**A**) Frequencies of response times (in ms) of all trials across all participants. (**B**) shows the percentage of positive feedback received for both conditions across all participants. The dark orange box represents the easy condition and light orange the difficult condition. (**Ci**) Multiple robust regression with behavioral adjustment from current to the consecutive trial (change in reaction times in ms) as outcome. Valence, expectedness, and their interaction significantly predict behavioral adjustment. To correct for multiple comparisons Bonferroni correction was used. (**Cii**) Unexpected negative feedback elicits the largest

(Figure 4-2 continued) behavioral adjustment. (**Di**) Frequencies of absolute behavioral adjustment from current to the consecutive trial (change in reaction times in ms) after negative (red colour bars) and positive feedback (blue colour bars). (**Dii**) Frequencies of behavioral adjustment (change in reaction time in ms) after unexpected (red) and expected (blue) feedback.

In ***Figure 4-2* D**, the absolute frequencies of averaged reaction time changes after negative/positive feedback (Di) and after expected/unexpected feedback (Dii) overall participants can be seen. Behavioral adjustments after negative feedback are bigger (M_{neg}= 230.39ms; M_{pos}= 145.06ms; $t(1982) = -42$, $p < .001$, d = 1.89) and within-subject more widely distributed (SD_{neg}= 180.20ms; SD_{pos}= 120.68ms; $t(1982) = -37$, $p < .001$, d = 1.68) than after positive feedback. Adjustments differ slightly with bigger adjustments and wider distributions within-subject for unexpected events (M_{unexp}= 197.52ms; M_{exp}= 180.53ms; $t(1982) = -8$, $p < .001$, d = 0.36; SD_{unexp}= 163.85ms; SD_{exp}= 152.02ms; $t(1982) = -7$, $p < .001$, d = 0.29). These results indicate that feedback valence and expectedness differentially affect subsequent behavior. To parse behavioral adjustments at finer levels of detail, we analyzed behavioral adjustments using multiple robust regression (***Figure 4-2* C**). Predictors included expectedness (expected vs. unexpected feedback), valence (positive vs. negative feedback), and the interaction between these factors. In addition, we included FCI and trial number as regressors of no interest. The results confirmed a main effect of expectedness (mean t = 0.10, $t(991) = 2.62$, $p < .05$ (bonf cor), d = 0.08, 95% CI [0.02, 0.17]) and valence (mean t = 3.89, $t(990) = 74.18$, $p < .001$ (bonf cor), d = 2.36, 95% CI [3.79, 4.00]; ***Figure 4-2* Ci**). Critically, the interaction between expectedness and valence showed that behavioral adaptations after negative feedback were larger, when negative feedback was less expected (mean t = -1.17, $t(990) = -18.65$, $p < .001$ (bonf cor), d = -0.59, 95% CI [-1.29, -1.05]; ***Figure 4-2* Cii**). This may suggest, that participants adjust their behavior more after negative feedback during broader target windows. In contrast, they adjusted their behavior after negative feedback less, when the target window was narrow (i.e., during the difficult task state). In other words, surprise had a reinforcing effect on behavioral adjustments when feedback was negative. We took this as motivation for further analyses trying to disentangle the differential influences of valence and expectedness on behavior and EEG correlates. As the absolute change in RT between two consecutive trials does not allow direct evaluation of the adaptivity of the behavioral adjustment, we calculated the variable RT_change as a

measure of performance adaptation for subsequent analyses investigating the association between feedback-related behavioral adaptions and EEG activity.

4.3.2 Electrophysiological Results

4.3.2.1 Event-related potentials

To gain a first impression of the EEG-data and to see, whether or not our paradigm evokes the expected ERP results, we show the averaged ERP waves for the crossed conditions, expected/negative, unexpected/negative, expected/positive, and unexpected/positive for two electrodes, FCz and Pz (***Figure 4-3***). At FCz (***Figure 4-3* Ai**), the amplitude during the time window of the FRN seems to be more pronounced when participants received negative (yellow and red) rather than positive feedback. When rewards were less expected (green line), the FRN signal was more positive than in the other condition (expected rewards, blue line), whereas the FRN on negative trials was similar for expected (yellow) and unexpected (red) feedback. The amplitude during the time window of the P3b component seems larger for unexpected events (green and red) than for expected events (blue and yellow). Visually comparing unexpected negative (red) and unexpected positive feedback (green) in the time window of the P3 complex, P3 amplitude seems to be affected by valence as well, as negative feedback increases the amplitude.

The EEG data in the time window of the P3 complex looks slightly different at electrode Pz (***Figure 4-3* Bi**) compared to FCz. The waveforms start do differ at around 200ms after feedback onset. Unexpected positive feedback (green) seems to be associated with the most positive signal from 220ms on. Nevertheless, the grand average ERP approach does not allow a temporally and spatially accurate determination of independent contributions to the neuronal signal: to investigate when and where which factors influence the neuronal signal and how they interact, it is essential to use other methods. The beta-value courses of the main predictors (***Figure 4-3* Aii; Bii**) are described in the context of the regression analysis (see 4.3.2.2).

The visual inspection of the grand average EEG activity supported our initial hypotheses concerning the influence of expectedness, valence, and their interaction on the feedback-locked EEG signal. For rigorous statistical testing and in order to disentangle this modulation of the signal temporally and spatially, and in dependence of

possible confounding factors, we applied a multiple robust single-trial regressions approach (Fischer et al., 2016; Fischer & Ullsperger, 2013).

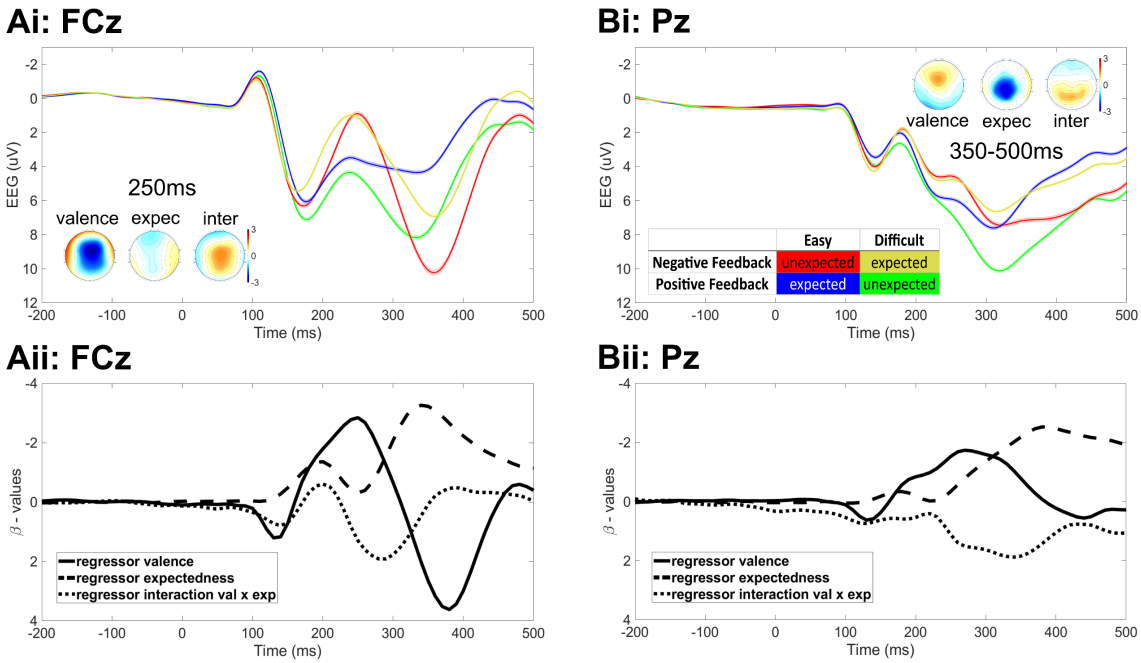

Figure 4-3. Grand average feedback-locked ERPs of the crossed conditions and beta-value courses of the main regressors in GLM 1a. (**Ai**) The averaged feedback-locked ERPs across participants at electrode FCz. The table in (**Bi**) shows the task design including the matrix of different task stages. Different colors represent the crossed conditions: expected negative feedback (yellow); unexpected negative (red); expected positive feedback (blue); unexpected positive feedback (green). Shades represent SEM (very small and therefore barely visible). Regression weight topographies of the regressors valence, expectedness, and their interaction are shown at 250ms after feedback. (**Aii**) The course of the beta-values from GLM1a of the regressors valence, expectedness, and their interaction at FCz are shown below. (**Bi**) The averaged feedback-locked ERPs across participants at electrode Pz. Regression weight topographies of the regressors valence, expectedness, and their interaction averaged across latencies from 350 to 500ms after feedback are shown. Regression weights in the topoplots are fdr (false discovery rate)-corrected and nonsignificant electrodes are masked in white. (**Bii**) The course of the beta-values from GLM1a of the regressors valence, expectedness, and their interaction at Pz are shown below.

4.3.2.2 Modulation of EEG signal by valence and expectedness of feedback

GLM 1a examined the effects of feedback *valence*, *expectedness*, and their *interaction* on neuronal EEG activity (***Figure 4-3* Aii, Bii** and ***Figure 4-4*** row **1-3**). From approximately 130ms until 300ms, there was a sustained negative effect of the *valence* regressor, that spanned over frontocentral electrode sites (***Figure 4-3* Aii**; ***Figure 4-4***, row **1**, shown from 200ms onwards). Thus, EEG amplitudes were more negative following negative feedback (most pronounced at FCz, peak @ 250ms, b = -2.63, *t*(991)

= -48.10, p = 1.81 x 10^{-261}, d = -1.52, 99% CI [-2.74 -2.53], *crit p* = 0.035), which indicates that valence coding is present in the typical latency range and topography of the FRN. Furthermore, there is another valence-specific but more parietal negative effect starting at 260ms until 340ms (***Figure 4-3* Bii**; ***Figure 4-4*** row **1**). This negative covariation was followed by a frontocentral positive covariation between 320 and 480ms (most pronounced at FCz, peak @ 380ms, b = 3.35, t(991) = 46.44, p = 6.47 x 10^{-251}, d = 1.48, 99% CI [3.21 3.49], *crit p* = 0.035; ***Figure 4-4***, row **1**). Topography and latency suggest that it is a major contributor to the P3a.

Within GLM 1a, we also tested the influence of *expectedness* on the EEG signal. This regressor shows a sustained and spatially broader fronto-centro-parietal effect on the P2/P3 complex between 240 and 530ms (most pronounced at Cz, peak @ 350ms, b = -3.26, t(991) = -60.32, p < 0.0001, d = -1.92, 99% CI [-3.37, -3.15], *crit p*: 0.0001; ***Figure 4-4***, row **2**). The signal is more positive for unexpected compared to expected outcomes. However, the effect of expectedness prior to 240ms seems to be less pronounced and more frontally distributed (Cz, smaller peak @ 190ms, b = -0.91, t(991) = -34.42, p = 1.93 x 10^{-171}, d = -1.09, 99% CI [-0.96, -0.86], *crit p* = 0.0001). Taken together, the FRN amplitude is influenced by valence, but is less affected by expectedness (between 200 and 300ms; $mean_{val}$ = -0.21 ± 0.004 vs. m_{exp} = -0.03 ± 0.003; t(991) = -41.46, p = 1.14 x 10^{-218}, d = 1.82). Valence *and* expectedness (i.e. global surprise, resp.) seem to influence the P2/P3 complex. Whereas the valence effect appears to be frontocentral, violation of expectedness shows a broader effect from frontocentral to parietal areas.

Additionally, we were interested in whether and how both regressors interact with each other as an interaction would be predicted for an RPE signal. Derived from the course of beta-values, from approximately 230ms on, the *valence* x *expectedness* regressor covaried positively with the neuronal activity (FCz, peak @ 280ms, b = 1.81, t(991) = 17.10, p = 1.28 x 10^{-57}, d = 0.54, 99% CI [1.60, 2.01], *crit p* = 0.032; ***Figure 4-4***, row **3**). By visual inspection, this broad frontocentral positive covariation extends to parietal areas over time. To summarize, the interaction of valence and expectedness shows a strong significant effect during the FRN latency range for both the early FRN-peak and for the later part of the FRN on fronto-centro-parietal electrodes. This underlines that the FRN encodes an RPE signal. To disentangle the interaction between valence and

expectedness, we conducted separate regression models by splitting the data in expected and unexpected trials (GLM 1b).

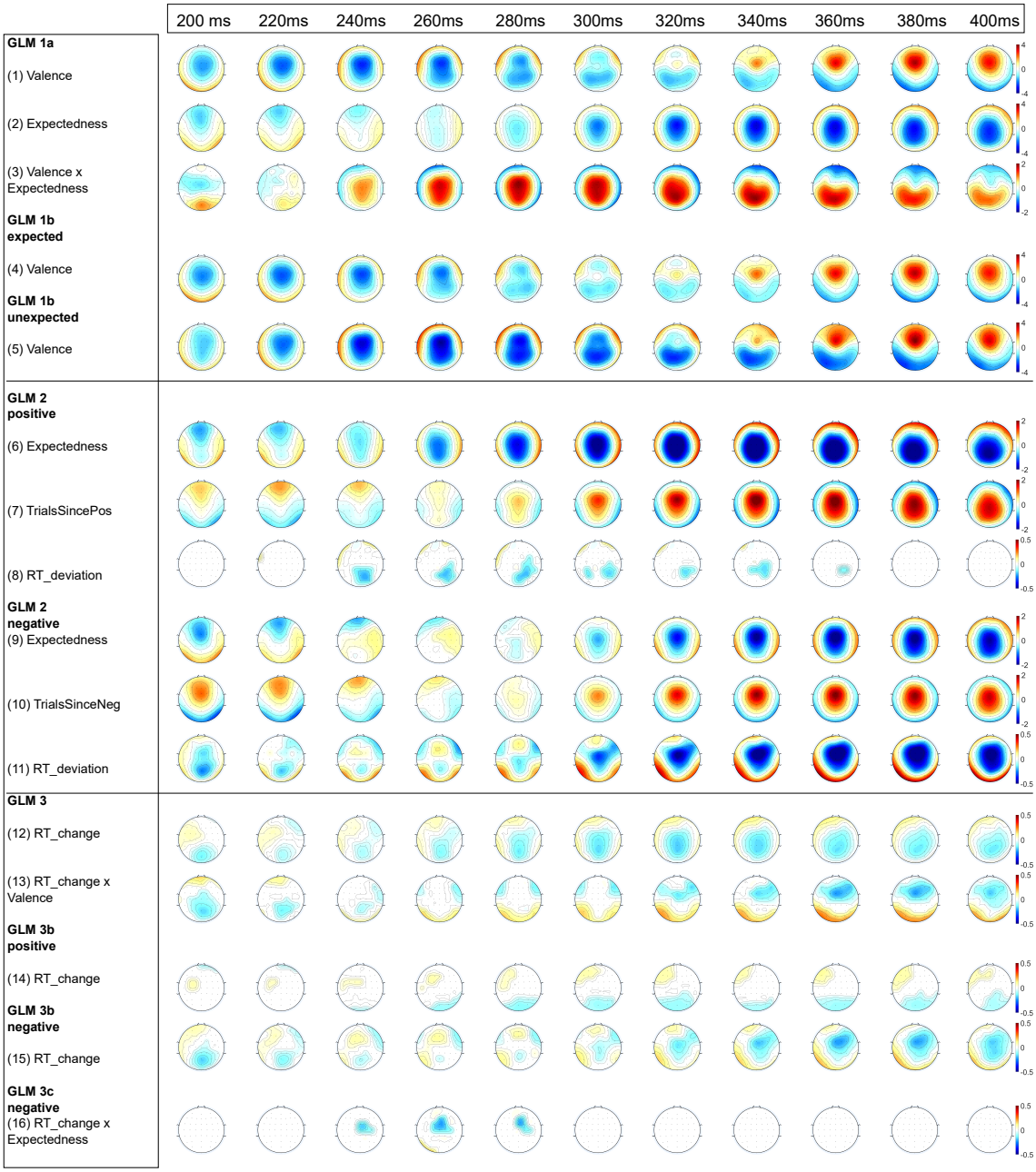

Figure 4-4. Multiple single-trial robust regression results for feedback-locked epochs. Feedback-locked regression weight (beta) topographies in 20ms time steps spanning from 200ms to 400ms are shown. The corresponding GLM and predictors are listed on the left. Blue colors are associated with negative covariations, red colors with positive covariations. Interpretation of the polarity depends on the coding of the predictor (see methods section 2.4.1). Scaling can be seen on the right and differs between regressors. For corrections of multiple comparisons, false discovery rate (FDR) was used. Nonsignificant ($p<.01$) data points are masked in white.

In GLM 1b, we found a negative covariation of *valence* with the neuronal signal from 160ms to 310ms in both, unexpected (FCz, peak @ 260ms, b = -3.55, $t(991) = -41.95$, $p = 6.64 \times 10^{-222}$, d = -1.33, 99% CI [-3.72, -3.39], *crit p* = 0.036; ***Figure 4-4***, row **5**) and expected trials (FCz, peak @ 240ms, b = -2.41, $t(991) = -37.89$, $p = 6.14 \times 10^{-195}$, d = -1.20, 99% CI [-2.53, -2.28], *crit p* = 0.033; ***Figure 4-4***, row **4**). Comparing the regression weights at their respective peaks, the valence effect during the time window of the FRN is more pronounced in unexpected trials ($m_{unexp} = -3.55 \pm 0.08$ vs. $m_{exp} = -2.41 \pm 0.06$; $t(991) = -11.49$, $p = 8.63 \times 10^{-29}$, d = 0.5). To formally test for spatiotemporal differences of valence coding between expected and unexpected feedback in the FRN time window, we extracted averaged b-values at FCz and Pz and at the latencies 200ms (± 20ms) and 260ms (± 20ms) from the valence regressor of the respective GLM 1b. These data were then analyzed in a three-way ANOVA with the factors latency, location, and expectedness. Results show a significant three-way interaction of expectedness x location x latency ($F(1,991) = 122.39$, $p < .001$, $\eta^2 = .11$) indicating that for expected events, valence is coded earlier (200ms) and more frontally (FCz), whereas for unexpected events, valence is processed later (260ms) and with parietal involvement (Pz).

Afterwards, until approximately 450ms, the signal covaries positively with *valence* in both, unexpected (FCz, peak @ 380ms, b = 3.57, $t(991) = 30.53$, $p = 8.39 \times 10^{-145}$, d = 0.97, 99% CI [3.34, 3.80], *crit p* = 0.036; ***Figure 4-4***, row **5**) and expected trials (FCz, peak @ 380ms, b = 3.27, $t(991) = 41.67$, $p = 4.33 \times 10^{-220}$, d = 1.32, 99% CI [3.11, 3.42], *crit p* = 0.033; ***Figure 4-4***, row **4**). Comparing the regression weights at their peak at 380ms, the effect is more pronounced in unexpected trials compared to expected trials ($m_{unexp} = 3.57 \pm 0.12$ vs. $m_{exp} = 3.27 \pm 0.08$; $t(991) = 2.40$, $p = 0.02$, d = 0.10). This result illustrates the involvement of surprise: if the event is unexpected, the positive-going valence effect on the P3b is stronger.

Taken together results from GLM 1b, aiming to disentangle the interaction of valence and expectedness of feedback, we found an effect of valence on the neuronal signal during the time windows of the FRN and the P3 for expected *and* unexpected events. In both, the time window of the FRN and the P3, the effect of feedback valence was more pronounced in unexpected events.

To *directly* compare the neuronal activity of expected and unexpected events in dependence of valence, we conducted GLM 2 on positive- and negative feedback trials separately. As can be seen in ***Figure 4-4***, row **6**, there was an influence of expectedness for positive-feedback trials at frontal and frontocentral electrodes starting before the FRN (90ms; FCz, peak @ 160ms, b = -0.56, $t(991) = -11.75$, $p = 6.10 \times 10^{-30}$, d = -0.37, 99% CI [-0.65, -0.47], *crit p* = 0.029; ***Figure 4-4***, row **6**) and spanning the entire FRN latency range. For negative-feedback trials, there is a small negative covariation of expectedness with the neuronal signal at frontal and frontocentral electrodes starting early around 160ms (FCz, peak @ 200ms, b = -0.86, $t(991) = -16.09$, $p = 6.47 \times 10^{-52}$, d = -0.51, 99% CI [-0.96, -0.75], *crit p* = 0.027; ***Figure 4-4***, row **9**). In contrast to the expectedness effect in positive-feedback trials, in negative-feedback trials there seems to be a weaker frontocentral effect in the second half of FRN latency ranges between 240ms and 280ms. In other words, for negative outcomes, the FRN is nearly unmodulated by expectedness. In contrast, for positive outcomes, the positive-going shift of the waveform is significantly pronounced when the outcome is unexpected. A stronger modulation on trials with positive feedback also applies to the later ERP components. During the time window of the P3, results show a significant negative covariation at central and centroparietal electrodes for the regressor expectedness for trials with positive feedback (Cz, peak @ 330ms, b = -2.49, $t(991) = -34.73$, $p = 1.51 \times 10^{-173}$, d = -1.10, 99% CI [-2.63, -2.35], *crit p* = 0.029; ***Figure 4-4***, row **6**), but also for trials with negative feedback (FCz, peak @ 360ms, b = -2.06, $t(991) = -22.04$, $p = 6.11 \times 10^{-88}$, d = -0.70, 99% CI [-2.25, -1.88], *crit p* = 0.027; ***Figure 4-4***, row **9**).

4.3.2.3 Modulation of EEG signal by local surprise

Next, we investigated neural coding of local surprise, that was reflected in two regressors within GLM 2, the *number of preceding trials since the last negative or positive trial appeared* and *reaction time deviation* from the target time of 1000ms. The results indicate that the longer it has been since the current feedback was last seen (i.e., the bigger the local surprise), the more positive the P3b. On frontal electrodes, the first predictor covaried positively with the neuronal signal in both positive- (Fz, @ 220ms, b = 0.62, $t(991) = 23.13$, $p = 5.57 \times 10^{-95}$, d = 0.73, 99% CI [0.56, 0.67], *crit p* = 0.031; ***Figure 4-4***, row **7**) and negative-feedback trials (Fz, @ 220ms, b = 0.82, $t(991) = 30.43$, $p = 3.81 \times 10^{-144}$, d = 0.97, 99% CI [0.77, 0.88], *crit p* = 0.033; ***Figure 4-4***, row **10**). This

effect is stronger in negative-feedback trials between 200 and 240ms (m_{pos} = 0.62 ± 0.03 vs. m_{neg} = 0.82 ± 0.03; $t(991) = -5.71$, $p = 1.46 \times 10^{-8}$, d = 0.25). From 280ms on, this predictor covaried positively with the neuronal signal in both positive- (FCz, peak @ 340ms, b = 2.07, $t(991) = 43.49$, $p = 5.48 \times 10^{-232}$, d = 1.38, 99% CI [1.97, 2.16], *crit p* = 0.031; ***Figure 4-4***, row **7**) and negative-feedback trials (FCz, peak @ 360ms, b = 1.96, $t(991) = 42.52$, $p = 1.19 \times 10^{-225}$, d = 1.35, 99% CI [1.87, 2.05], *crit p* = 0.033; ***Figure 4-4***, row **10**).

For the predictor *reaction time deviation* from the target time of 1000ms, there is a negative covariation within the time frame of P3 only in negative trials (FCz, peak @ 370ms, b = -0.67, $t(991) = -24.60$, $p = 1.19 \times 10^{-104}$, d = -0.78, 99% CI [-0.72, -0.62], *crit p* = 0.023; ***Figure 4-4***, row **11**). If the participant was close to 1sec, but negative feedback followed, subjective surprise might be enhanced and in consequence, P3 is more positive. This does not apply for the positive prediction error, considering only a small parietal to occipital activation can be found in positive-feedback trials between 240 and 340 ms (***Figure 4-4***, row **8**). In conclusion, the present data (GLM 2) indicate that the P3 complex is driven by global *and* local surprise.

In GLM 2, we found that expectedness of feedback has only a small effect on the neuronal signal during the time window of the FRN. Meanwhile, global surprise during the P3 time window influences the neuronal signal in positive-feedback and negative-feedback trials. Furthermore, we found evidence for a positive covariation of local surprise and the P3.

4.3.2.4 Adaptation

We were interested in how the EEG signal, and especially, the FRN is associated with behavioral adaptations after making a false response. Therefore, we implemented another behavioral predictor in GLM 3, *reaction time change* (*RT_change*). It represents feedback adaptation in the consecutive trial, while positive values imply improvement (i.e. getting closer to 1sec) and negative values decline (further away from 1sec) in performance. Furthermore, we included the interaction of valence and adaptation, because we assume a stronger influence of adaptation after negative-feedback trials. Surprisingly, we only see a very small association between *RT_change* and the neuronal activity from 130ms on (CPz, peak @ 220ms, b = -0.04, $t(991) = -3.39$, $p = 0.001$, d = -0.11, 99% CI [-0.06, -0.02], *crit p* = 0.047; ***Figure 4-4***, row **12**). However, there is a

negative covariation between *RT_change* at centroparietal electrodes (CPz, peak @ 320ms, b = -0.17, $t(991) = -12.61$, $p = 6.36 \times 10^{-34}$, d = -0.40, 99% CI [-0.19, -0.14], *crit p* = 0.047; ***Figure 4-4***, row **12**) and the interaction of *feedback valence* x *RT_change* at frontal electrodes with the neuronal signal around 350 ms after feedback (FCz, peak @ 370ms, b = -0.22, $t(991) = -7.58$, $p = 7.80 \times 10^{-14}$, d = -0.24, 99% CI [-0.28, -0.17], *crit p* = 0.008; ***Figure 4-4***, row **13**). These counterintuitive findings imply that an improvement in the consecutive trial is associated with a smaller P3a amplitude. To disentangle the effects of the significant interaction between valence and behavioral adaptation, we split the data into negative- and positive-feedback-trials and used RT_change as a regressor of interest (GLM 3b). Within positive-feedback trials, RT_change does not seem to have a systematic influence on neuronal signals contributing to the FRN or P3 complex (***Figure 4-4***, row **14**). For negative trials, there is a small negative covariation of RT_change with the neuronal signal at parietal electrodes in the time frame of the early FRN (Pz, @ 200ms, b = -0.15, $t(991) = -10.88$, $p = 4.18 \times 10^{-26}$, d = -0.35, 99% CI [-0.18, -0.13], *crit p* = 0.011; ***Figure 4-4***, row **15**). Additionally, results show a negative covariation of RT_change with the neuronal signal in the time window of the P3 for negative-feedback trials (FC2, peak @ 360ms, b = -0.21, $t(991) = -13.00$, $p = 8.59 \times 10^{-36}$, d = -0.41, 99% CI [-0.24, -0.18], *crit p* = 0.011; ***Figure 4-4***, row **15**). Because participants adjusted their behavior depending on the expectancy of negative feedback, we investigated the influence of the interaction of expectedness and RT_change after negative feedback on the neuronal signal in GLM 3c. Indeed, a frontoparietal negative covariation between 240ms and 280ms occurs in negative-feedback trials (FCz, peak @ 260ms, b = -0.19, $t(991) = -4.99$, $p = 7.14 \times 10^{-7}$, d = -0.16, 99% CI [-0.27, -0.12], *crit p* = 0.0004; ***Figure 4-4***, row **16**).

4.4 Discussion

4.4.1 Factors influencing the early feedback-related neuronal signal

The current study found, based on a large sample of 992 participants, that valence and expectedness influence the neuronal signal and behavioral adjustments after feedback. ***Figure 4-5*** summarizes the contributions of feedback-related performance monitoring components to feedback-locked EEG dynamics in a schematic sketch. As can be seen, the effect of valence starts early around 180ms at frontoparietal areas. Another negative covariation joins in from 260ms on at centroparietal electrodes. This second, more parietal valence-specific effect was already reported in other studies (Gentsch et al., 2009; Ullsperger, Danielmeier, & Jocham, 2014) and appeared independently from frontocentral parts (Gentsch et al., 2009). Since the classification of the signal is still unclear, future research should address the origin of this phenomenon. With respect to global surprise or expectedness, we see a weaker frontal effect starting at an early latency even before 180ms (***Figure 4-5***). The interaction of valence and surprise evokes a broad and sustained frontocentral effect from 230ms on and it extends to parietal areas over time.

Negative feedback was associated with a sustained negativity in the FRN latency range. In the same time window, we found a negative covariation of the feedback-locked EEG activity for the expectedness regressor. These findings go in line with previous research (Holroyd & Krigolson, 2007) and support the RL-account in a way, that the FRN is not only driven by mere surprise. Moreover, we found an interaction of expectedness and valence affecting the EEG signal in the latency range of the FRN. This result supports the RL theory that the FRN is affected by a signed RPE signal. The FRN modulation by expectedness was stronger after positive feedback than after negative feedback. Hence, one could also interpret the present findings in the context of a RewP. Since this valence-specific effect occurs only in interaction with expectedness, this interpretation does not seem straight forward to us. In other words, EEG dynamics in the time range of the FRN are more modulated by outcome expectancy after positive compared to negative feedback. RPEs are therefore stronger represented and have a larger "dynamic range" during positive events than during negative events. In this case,

the RPE appears to be driven by surprise. In contrast, when the RPE is driven by valence, we found a stronger, later, and more parietal modulation of the FRN by valence for unexpected events than for expected events. This finding rather supports recent evidence suggesting that the FRN reflects some combination of reward- and salience-prediction error encoding (Glazer & Nusslock, 2021). In line with that, new studies state that the signal in the latency range of the FRN reflects independent processing of both better-than-expected and worse-than-expected outcomes (Bernat et al., 2015; Bernat et al., 2008; Foti et al., 2015; Hoy et al., 2021). The FRN is proposed to reflect the neural response to a negative RPE (losses) associated with theta band oscillatory perturbations (Cavanagh et al., 2012; Cavanagh & Frank, 2014; Cavanagh et al., 2010; Cavanagh & Shackman, 2015), whereas the RewP is proposed to reflect the neural response to a positive RPE (wins) associated with delta band activities (Cavanagh, 2015). Unfortunately, in the present study, there are some limitations that restrict the interpretation of the RPE, particularly the differentiation into positive and negative RPEs. Due to the task design, the probability and valence of the feedback are intertwined.

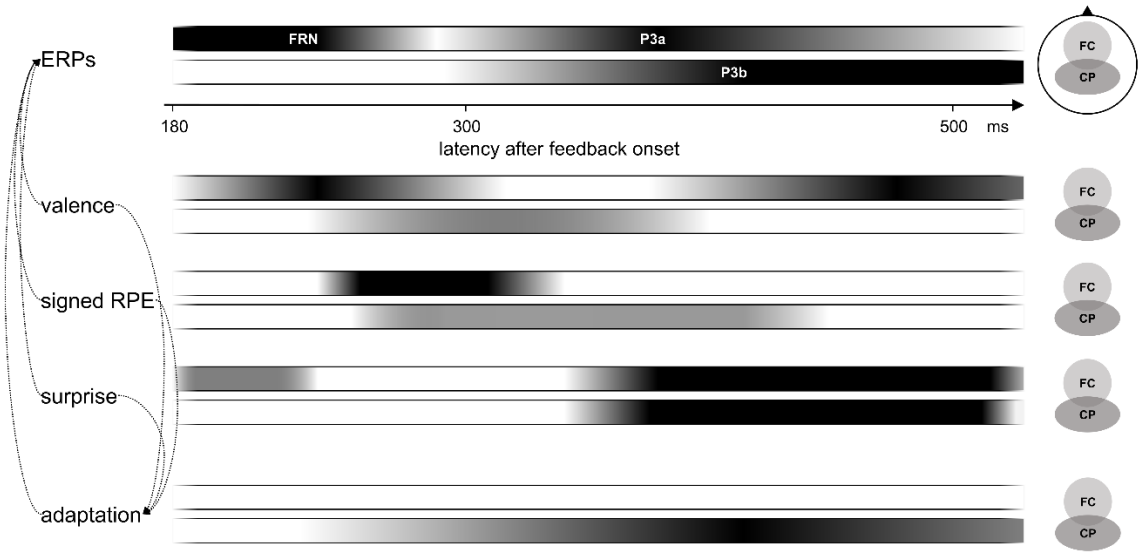

Figure 4-5. Schematic representation of influencing factors and their manifestation in the neuronal signal and ERPs after feedback. ERP = event-related potential; FC = frontocentral; CP = centroparietal.

4.4.2 Factors influencing the later feedback-related neuronal signal

Results for the neuronal signal at later latencies show also an influence of expectedness and valence. Regarding surprise, there is a sustained centroparietal effect

from around 300ms on (***Figure 4-5***). From 340ms after feedback, a valence-specific frontoparietal positive covariation appears and finds its peak at 380ms. Concerning the P3 complex, the influence of global surprise could be replicated (de Bruijn et al., 2004; Fischer & Ullsperger, 2013; Mars et al., 2008): we showed a dependence of the neuronal signal during the time window of the P3b at centroparietal areas on expectedness. Valence elicited a frontoparietal effect from 340ms on. This P3a-pattern could therefore represent a stimulus-driven attention mechanism. Furthermore, the P3 complex was most positive for unexpected positive-feedback outcomes. This finding is consistent with the results of Hajcak et al. (2007) and Severo et al. (2018). Walentowska et al. (2019) had similar findings: the P3b amplitude was larger for positive than negative feedback in unexpected events. Because surprising positive events can be very helpful in increasing the performance, this phenomenon could be interpreted as a goal-relevant action-value-updating of this specific feedback. Additionally, we addressed surprise on a more trial-by-trial basis and examined the impact of a locally surprising event. In line with previous evidence (Mars et al., 2008; Nieuwenhuis et al., 2005; Squires et al., 1976), results show a dependence of the P3 complex to local surprise in a way, that the P3 is more positive when local deviants appear.

4.4.3 Link between FRN and behavioral adaptation

Concerning behavioral adaptation, an association between the size of the FRN and the amount of behavioral adjustment in the following trial in combination with feedback valence was assumed: after negative feedback, a stronger effect on the neuronal signal in the time window of the FRN is expected to be associated with a bigger change in reaction time than after positive feedback. In a more exploratory way, we created a measure for behavioral adaptation that also depicts performance improvement and decline. Results show only a small association between adaptation after feedback and the neuronal signal during the time window of the FRN both overall trials and for trials with negative feedback, mostly at parietal electrodes. Previous research was able to link the FRN to behavioral adaptations depending on feedback valence and expectedness (Arbel et al., 2013; Cavanagh, 2015; Holroyd & Krigolson, 2007). One reason why we could not reproduce these findings is the nature of the feedback used in the present task. The feedback contained information about the correctness of the response but did not tell

the participants in which direction to adjust their behavior in order to improve their performance in the upcoming trial. At the same time, our measure of behavioral adaptation did contain directional information about whether one is getting closer or further away from the correct response, and thus did not correspond to the information content of the feedback itself. In order to use feedback profitably, the information content is crucial. It has already been shown that this also affects the neuronal signal (Cockburn & Holroyd, 2018). Future experiments should therefore work with differentially informative feedback and corresponding outcome measures to more specifically capture the relationship between the feedback-processing neuronal signals and behavioral consequences. The unspecific association between behavioral adaptations and feedback-related EEG activity in our task might reflect participants' attempt to adjust their behavior in an exploratory manner following negative feedback. The dependence of behavioral adaptation on both feedback valence and expectedness might suggest that participants infer that it is adaptive to adjust their behavior more after unexpected negative feedback. In this task-state, the target time window is wider, hence larger adjustments are more likely to improve behavior. Because participants are not told whether their time estimation was too fast or too slow, they are forced to randomly adjust their behavior after negative feedback and try to memorize their time estimation after positive feedback. This process could be reflected in the sustained parietal positive response seen after positive feedback. For the present study, however, this assumption has to remain a tentative explanation.

In line with the findings mentioned above (4.1), adaptive behavior and learning after feedback can have different functions depending on the type of RPE: an unsigned PE may enhance selective attention based on known task rules (Danielmeier et al., 2015; Danielmeier et al., 2011; King et al., 2010), while a signed PE determines the direction of reinforcement learning, that is, to repeat or avoid an action (Ullsperger, Fischer, et al., 2014). Arbel et al. (2013) showed that the FRN elicited by negative feedback was not correlated with long-term learning outcomes, whereas positive-feedback-associated FRN was correlated with the learning outcomes. In contrast, Cavanagh (2015) hypothesized a differentiation in hierarchically distinct types of prediction error, where delta band activity linked to a rewarding event motivates immediate behaviors, while theta band activity linked to punishing events initiates long-term behavioral adjustments

(Cavanagh & Shackman, 2015). In the present study, we demonstrate a small effect of the interaction of consecutive behavioral adaptation with outcome expectedness on the neuronal signal between 240 and 280ms for trials with negative feedback. This may provide evidence for the connection between negative RPE and behavioral adjustments. Nevertheless, the FRN varies due to different feedback characteristics and the amount of feedback information (Cockburn & Holroyd, 2018) and is therefore highly task dependent. This could be one reason for inconsistencies in existing evidence. Studies finding inter-individual differences (van Noordt & Segalowitz, 2012) or studies related to mental disorders (Endrass et al., 2013; Keren et al., 2018; Webb et al., 2017) also show that the FRN does not reflect a monolithic block with only one unique functional interpretation. Rather, the task, context, etc., must be considered to characterize the factors that contribute to the FRN. These contributions and their representations may vary independently (Stewardson & Sambrook, 2021), leading to variations in ERPs. This can even lead to latency shifts in the individual effects, making it difficult to unambiguously assign FRN and P3. Therefore, the task should always be considered when interpreting different findings concerning the FRN and adaptive behavior.

4.4.4 Link between P3 and behavioral adaptation

Since the feedback-related P3 had been associated with behavioral adaptation (Fischer & Ullsperger, 2013; Jepma et al., 2018; Jepma et al., 2016), the P3 amplitude could be expected to correlate positively with changes in behavioral performance after feedback. Surprisingly, there was a contrarian association between behavioral adaptation and the neuronal signal in the P3 time frame at frontoparietal areas (***Figure 4-5***): the neuronal signal appears less positive when behavioral adaptation to the next trial is greater, i.e. that performance has improved. Nassar et al. (2019) and Kirschner et al. (2022) predicted adjustments in behavior based on the amplitude of the P3, but as a function of the source of surprise. If the surprise was uninformative, the P3 negatively predicted learning. As mentioned earlier, the feedback in the present study was uninformative in a way that it was not directional. Thus, it can be concluded that the counterintuitive association between P3 and behavioral adjustment in the present study may reflect the negative prediction from above (Kirschner et al., 2022; Nassar et al., 2019). Additionally, Cavanagh (2015) found an association between the P3 complex and

behavioral adaptation: delta band phase dynamics observed in the P3 appear to be involved in strategic behavioral adjustments like the degree of response time speeding. Interpretation of the authors suggests that the processes underlying FRN and P3 reflect hierarchically different levels of prediction errors, reward and state prediction errors, respectively. Whereas reward prediction errors give information to a model-free learner on a trial-and-error basis, state prediction errors inform model-based systems by more complex forward predictions. If the P3 is modulated by state prediction errors and therefore depends on an agents' decision-making policy, a richer learning environment may be needed to reveal the relationship between behavioral adaptation and the neuronal signal during the corresponding time frame (like Chase et al., 2011).

Taken together, the investigated ERP deflections are not only due to single influencing factors, but are complex representations composed of several influencing components. As ***Figure 4-5*** clearly shows, even within FRN latency ranges there are temporally and spatially distinguishable effects that are difficult to characterize using an ERP averaging approach. Also, previous methodological considerations (Glazer et al., 2018; Williams et al., 2021) indicate the difficulty to isolate the FRN and confounds arising from component overlap. Therefore, they underline the importance of teasing apart the stages of feedback processing to integrate individual reward-related ERPs in a more holistic view and to capture the broader temporal dynamics (Foti & Weinberg, 2018; Glazer et al., 2018). There are already some attempts to do so (Gheza et al., 2018; Sambrook & Goslin, 2016). Moreover, time-frequency analysis could help future research identify and differentiate cognitive processes underlying different types of RPE and their roles in behavioral adjustments to improve performance.

4.4.5 Limitations

Methodological limitations include that feedback-related EEG dynamics may have been modulated by offset-related visual ERPs after 350ms past feedback, but it is unlikely to affect our results with respect to the regressors of interest. Additionally, previous findings have shown an influence of valence effects on the amplitude of the FRN depending on the perceptual salience of the feedback stimuli (Y. Liu et al., 2014). In the present study, we cannot rule out this effect because feedback color was not counterbalanced across participants.

4.4.6 Conclusion

In the present study, we disentangled the functional relevance of independent contributions to electrophysiological correlates of feedback processing in a big sample of N=992 participants by using a novel approach in EEG analysis and a time estimation paradigm. While we have previously used instrumental learning tasks (Burnside et al., 2019; Fischer & Ullsperger, 2013; Kirschner et al., 2022), we were able to replicate typical feedback-related components with the present task despite it involves less opportunities to adapt behavior. The results of the present study support the view that the FRN is driven by a signed RPE and furthermore influenced by global surprise. Depending on what drives the RPE most –expected value or obtained outcome–, different modulations of the FRN are possible. Whereas the FRN was less influenced by global and local surprise, unexpected events elicited a larger P3b. An association between behavioral adjustments and the P3 might indicate a representation of different RPE levels and types within the components involved in feedback processing. With the help of a big sample size and a regression approach that allows the simultaneous investigation of multiple independent variables, we obtain information on the temporal and spatial variance of the contributing effects on the neuronal signal.

5 General Discussion

The presented work aimed to enhance our understanding of human foraging decision-making mechanisms and feedback processing. It focuses on investigating how people make foraging-decisions in a patch-leaving environment, their adaptation to different reward rates, and their performance compared to an optimal MVT-inspired policy. Additionally, the studies explore potential neural correlates associated with continuous foraging-decisions and how they manifest. Regarding feedback processing, this research studied the type of RPE reflected in the neuronal signal during the latency of the FRN. Furthermore, it investigated factors influencing the latency of the P3, specifically the role of unexpectedness and valence. I also explored whether the neural signal during the FRN and P3 is associated with adaptive behavior after feedback.

To achieve these objectives, two experiments were conducted, recording EEG data while participants performed a foraging-decision task and a feedback-related paradigm. EEG was chosen for its high temporal resolution. Both studies utilized a robust single-trial regression approach to analyze independent contribution of task manipulations to the EEG signal and avoid reliance on specific ERP quantifications, allowing for a descriptive examination of the effects.

As hypothesized, participants in study I adapt their patch-leaving decisions in the direction an MVT-inspired optimal agent suggests, but not to a sufficient degree. The adaptation to the FRR was closer to the optimum than the adaptation to the BRR. Simultaneously, participants were closer to the optimum in a rich than in a poor environment, while generally staying shorter than optimal. On a neuronal level, results show a sustained impact of activity in the delta frequency range reflecting the BRR, while decreases in beta power before the response to leave is executed, reflect changes in local reward rates. Activity in the theta frequency range appears to represent a common component associated with both environmental changes as well as foreground reward rates.

Regarding the processing of feedback, the results provide evidence supporting the notion that the FRN is influenced by a signed RPE. Depending on the predominant driver of the RPE – whether it is the expected value or the obtained outcome – various

modulations in the time frame of the FRN can be observed. Additionally, the P3 complex is influenced by both surprise and valence. Interestingly, at the FRN level, only weak correlations with behavioral adaptations were found. Conversely, a negative correlation was found for the P3, suggesting that this component may be more closely associated with certain behavioral changes in response to feedback.

In light of these findings, I will discuss implications, the connection between decision-making and feedback processing, and speculate about future perspectives in the following sections.

5.1 Human foraging behavior compared to a normative agent

The ability to make foraging-decisions empowers individuals to navigate complex decision scenarios and make choices that optimize their desired outcomes, even in situations where future prospects are uncertain. Humans are often expected to behave optimally, efficiently utilizing time to maximize output while minimizing energy expenditure. Neuroscientific research aims to construct models to understand and replicate human decision-making. However, when comparing these models to actual human behavior, it is evident that individuals approximate optimal adaptive behavior but often deviate from the predicted optimum. One possible reason for this deviation is that the computational models involve complex and effortful calculations, which may not align with the brain's preference for efficiency and energy conservation. Humans may rely on simpler heuristics, leading to solid but less than perfect solutions that adhere to the principle of cost-benefit trade-offs (Findling et al., 2021). In recent research, a combined approach that integrates RFL models with complementary parameters or heuristic mechanisms provide the most accurate fit to human behavior in various exploration-exploitation scenarios (Brands et al., 2023; Cogliati Dezza et al., 2017). Additionally, the assumptions made within the computational models do not reflect the reality of individuals and often overlook human biases. For example, the phenomenon of overharvesting in foraging paradigms, as suggested by Harhen and Bornstein (2023), could result from a lack of complete environmental knowledge, leading humans to work with available information and experiences at a given time. This scenario may also apply

to our current Study I: the model assumes a consistent average reward rate within a single environment. However, especially initially, the environment is highly uncertain to participants. Consequently, the representation of the environment's average rate adapts dynamically with each encountered patch. We also incorporated a model accommodating this dynamic adaptation. Interestingly, in this specific instance, the predictions of both models exhibited minimal divergence, leading us to favor the simpler model. Other research (Acuña & Schrater, 2010; Busemeyer et al., 2019; Kilpatrick et al., 2021; Sims et al., 2013) has indicated that accounting for this dynamic evolution, e.g., using Bayesian updating, demonstrates that individuals' foraging behavior is not irrational but instead reflects optimal choices under uncertainty.

The main goal of this work, however, was not to find the best model to explain participants' behavior but rather to create a normative agent inspired by the widely used MVT approach to compare predicted and observed results and correlate them with the neuronal signal. By doing so, study I aimed to reveal neural correlates of optimal or suboptimal behavior. The initial hypotheses regarding the behavioral data were largely confirmed, revealing independent adaptations to variations in the FRR and BRR in the direction suggested by the MVT. However, this adaptation was not sufficient, and participants tended to adjust their behavior less effectively to higher-quality patches. Higher-quality patches in foraging scenarios are associated with longer stay/leave times, implying a more delayed reward. The bias towards leaving higher-quality patches prematurely could be attributed to temporal discounting, where distant rewards are devalued, leading individuals to prioritize smaller but more immediate rewards (Odum, 2011; Peters & Büchel, 2009). Consequently, participants may exhibit a preference for leaving high-quality patches too early, thereby potentially missing out on even greater rewards that could have been obtained with more extended stays. Moreover, participants performed worse when they were situated in the poor compared to the rich environment. Their behavior deviated significantly from the optimum in the poor environment, indicating challenges in optimizing decisions under unfavorable conditions. The phenomenon of an optimism bias, in which positive PEs are more heavily weighted (Sharot et al., 2012), could make it more difficult to adapt to resource-poor environments that are more likely associated with negative PEs. This bias towards favoring positive outcomes could prevent individuals from accurately assessing the true

costs and risks associated with foraging decisions in challenging or less rewarding environments. As a result, individuals may be less inclined to adjust their behavior appropriately. A surprising finding emerged, contrary to what the literature might have assumed: across conditions, participants actually stayed for shorter durations than optimal. As discussed earlier, this might be a paradigm-specific effect: A sense of urgency is created by emphasizing time costs, which leads to leaving options early to save time. These results shed light on the complexities of human decision-making during foraging tasks, highlighting the challenges individuals face in achieving optimal strategies, especially in changing and unfavorable environments. Both the computational models and observed behavior emphasize the importance of an agent's current foreground i.e. currently experienced reward rates (FRR) being compared to the longer running or assumed BRR.

Returning to the model introduced at the beginning of this thesis, foraging decisions involve the stages valuation and decision-making based on Rangel et al.'s (2008) cycle of decision-making. Within this framework, the values to compare are the value of an option encountered (to decide whether to engage with it) and the average value of the environment (to determine whether the environment is sufficiently rich to forego the encountered option and continue searching for better alternatives elsewhere). For a model of sequential decision-making with continuous harvest phases, these two stages need to loop multiple times within an encountered option until the decision to leave is made, allowing for an evaluation of this action. To evaluate the entire environment, one would need to loop through the entire process multiple times to estimate its overall quality. This assumption suggests that there are different mechanisms guiding decision-making between options within an environment and between different environments themselves. The complexity of this task raises the question of whether there is a specific, or perhaps multiple, neural mechanism optimized to solve this unique and ecologically significant challenge. The next section will address the contribution that this thesis can make to answering this question.

5.2 Neuronal dynamics of the human foraging process

The concurrently assessed electrophysiological data yielded significant results, providing neural representations of participants' choice strategies for adjustments to

both BRR and FRR during a foraging task. As expected, theta and beta activity play a role in the decision to leave a patch within a foraging scenario. Moreover, changes in the frequency range of delta appear to be associated with the specific environment in which participants are engaged. This sustained impact is observed from the beginning of the harvest phase until the decision to leave is executed. Therefore, delta oscillations might track broader contextual changes in the environment. Furthermore, reductions in midfrontal theta power are indicative of a shared component that tracks trial-by-trial variations in FRR and environmental changes. On the other hand, time-frequency analyses showed that decreases in beta activity are linked to variations in FRR. The results suggest that both beta and theta activity might inhibit the execution of the decision to leave until a certain threshold is met. It seems that participants delay their decision to leave until the representation of a current reward rate equals the average reward rate of the environment. These findings suggest that the neural representation of a foraging decision as an ongoing process, initiated at the start of a harvest phase, ramping up to a specific threshold, and ultimately culminating in the execution of the leaving decision when that threshold is met. The neural activity observed throughout the foraging process highlights the complexity and dynamic of decision-making during foraging tasks, providing valuable insights into the underlying mechanisms that guide our choices in such adaptive scenarios. The present work is the first utilizing electrophysiological data to illustrate the temporal dynamics of foraging decisions. However, many questions remain unanswered. Therefore, this thesis should motivate future work to determine the robustness and comprehensiveness of these findings. Given the high temporal resolution of EEG, it proves to be a suitable method to map this complex interplay of different processes. The interaction of different frequency activities proves insightful in illustrating these dynamics. It has also been suggested that certain frequencies, such as theta, are responsible for specific processes and coupling with another frequency, like gamma, which 'sits on top' and is possibly synchronized with the phase of theta, reflects the unique decision mechanism that distinguishes between exploiting and exploring (Domenech et al., 2020).

Due to previous literature linking the foraging process with the aMCC (Kolling et al., 2012; Wittmann et al., 2016), our analyses focused on frontocentral and central topographies. As expected, we found frontocentral electrophysiological dynamics,

supporting existing evidence. However, a surprising discovery was the presence of effects on electrodes in a more parietal position of the scalp. While the low spatial resolution of EEG could potentially account for these distortions, it is also plausible that these results indicate the involvement of a broader neural network cooperating during the process of making foraging decisions. In addition, the occurrence of activity in the delta frequency range argues for this, as oscillations in the low frequency range are commonly associated with coordination of larger-scale networks. The encoding of the leave decision was reflected in large-scale oscillatory power changes, possibly controlling a network of entire circuits responsible for context changes, since monitoring reward rates is relevant to many different decision processes such as vigour and willingness to exert effort. Findings of Laureiro-Martínez et al. (2015) support this hypothesis by discussing the involvement of various known brain circuits in the exploration-exploitation dilemma. Future research should delve into this indication further and explore the collaboration of the aMCC with other brain regions like the vmPFC to unravel the complexity of this decision-making process. By investigating the interactions and functional connectivity between various brain regions, we can gain a deeper understanding of how different neural networks work together to facilitate optimal decision-making during foraging tasks.

In the present study, participants were aware of the type of environment they were in (rich or poor). However, the precise representation of the current environmental value remains uncertain. Referring to the discussion on the operationalization of the average reward rate (see 1.1.2), understanding how the value of the environment is conceptually and neuronally reflected, as well as how possible future encounters are represented, is crucial. Addressing this question, Wittmann et al. (2016) made partial strides by modeling an expectation of prediction errors and linking it to activity in the anterior midcingulate cortex (aMCC). Their findings revealed that reward representations in the aMCC played a predictive role in individuals' decisions for exploitative or explorative behavior. While this provides valuable insights, further research is warranted to refine the neuronal representation of environmental value.

5.3 Independent contributions to neural correlates of feedback processing

The results of study II revealed independent effects of valence and expectedness on the electrophysiological signal following feedback. Moreover, the interaction of valence and expectedness influenced the signal in the latency range of the FRN, supporting the notion that the FRN reflects a signed RPE. In this respect, the hypothesis can be confirmed. Interestingly, depending on what drives the RPE – salience or reward – distinct effects were observed on the EEG signal, suggesting a combination of reward- and salience-based prediction error encoding in the FRN. This aligns with recent evidence proposing that the FRN measures the difference between two independent cognitive processes (Bernat et al., 2015; Foti et al., 2015; Hoy et al., 2021). Time-frequency analyses have garnered increasing interest in the field of feedback processing, offering valuable insights into the neural correlates of positive and negative prediction errors (Bernat et al., 2011; Cavanagh et al., 2010; Zheng & Mei, 2023). Delta activity is thought to reflect the processing of rewarding events to motivate related behavioral adjustments (Cavanagh, 2015), while frontal theta activity is associated with neural signals of surprise and uncertainty (Cavanagh et al., 2012; Cavanagh & Frank, 2014). As a next step, spatiotemporal processes are worth exploring further to gain a comprehensive understanding of feedback-related neural dynamics.

Furthermore, the study successfully replicated previous evidence linking the P3 complex to surprise, and results indicated an effect of valence on the signal during the latency range of the P3: a more positive signal when the feedback was unexpected and positive. Thereby the second hypotheses of this study can be confirmed.

Regarding learning from feedback and its neural correlates, the initial expectation was to find an association with the EEG signal in the latency range of the FRN. However, the results did not confirm this hypothesis, as the association in the FRN range was found to be weak. Instead, a less pronounced signal in the latency range of the P3 was observed in conjunction with greater behavioral adaptation in the upcoming trial and improved performance. This suggests that the P3 may play a more prominent role in the learning and adaptation processes following feedback, compared to the FRN. Contrarily,

Yan et al. (2023) found that the RewP amplitude predicted timing behavior for the upcoming trial. The relationship between the P3 and delta activity, which is associated with the motivation of action selection (Cavanagh, 2015), remains to be clarified. Additionally, while Cavanagh et al. (2010) linked also frontal theta activity to behavioral adaptation, the evidence on neural correlates of learning from feedback is still sparse. Consequently, more studies are required to gain a better understanding of the underlying electrophysiological processes and how individuals adapt their behavior based on RPEs.

Ultimately, it is essential to acknowledge that results on neural mechanisms for feedback processing heavily depend on the quantification of feedback-related signals and task context. To gain a comprehensive understanding of these processes, it is crucial to thoroughly investigate these influences, and fortunately, some studies addressing these aspects. For instance, Williams et al. (2021) conducted a comparison of different quantification methods and recommended to consider both ERPs and time-frequency analyses to provide a comprehensive picture. Furthermore, various studies explored the impact of contextual factors on the EEG signal during feedback processing, such as reward magnitude (e.g., Bellebaum et al., 2010), feedback delay (e.g., Weismüller & Bellebaum, 2016), motivation (e.g., Overmeyer et al., 2023), learning (e.g., Bellebaum & Daum, 2008), and memory (e.g., Albrecht et al., 2023). However, despite the progress made, many questions in this field remain unanswered. To facilitate a more comprehensive understanding and integration of findings, systematizations in the form of meta-analyses are necessary. Such meta-analyses can help identify common trends, establish robust patterns, and reveal potential discrepancies across studies. Additionally, examining and comparing data from different tasks will further enhance our knowledge of the generalizability and reliability of neural mechanisms underlying feedback processing.

5.4 The cycle of foraging decision-making and feedback processing

Revisiting the cycle of decision-making as proposed by Rangel et al. (2008), we can now explore how this process incorporates foraging decisions and connect it with the

findings of the present work. The evidence from our and other studies suggests that the foraging decision-making process involves some form of evidence accumulation during the foraging phase, beginning early during harvest and leading up to a threshold that triggers a specific action and outcome. This outcome is then compared to an expectation, and prediction errors are calculated, aligning with the principles of the RFL model. In Wittmann et al.'s (2016) study, the average reward rate is considered as the long-term average value of the currently exploited option instead of the value of concurrently alternative options. Comparing this average to the recent value of the option yields a measure of reward rate change, a PE. Since we already know quite a bit about how PEs are used as feedback and learning signals, this provides an opportunity to investigate both foraging decision-making and feedback processing together and uncover their interrelationships. A close linkage of these two processes is also evident from the fact that both foraging decisions and the need for behavioral adaptation are represented in aMCC (Clairis & Lopez-Persem, 2023). In our foraging paradigm, participants receive feedback after making a decision to leave, presenting an avenue for further exploration. Future research could investigate how reward rates, especially the average reward rate of the environment as a value representation, are updated during the feedback phase. During value-based decision-making, specific expectations are formed, which later influence PEs and the updating of reward rate representations. Exploring whether the knowledge of the value of options and related value-based choices between distinct options may predict the decision to leave could offer valuable insights for a cycle of sequential decision-making. Hassall and Krigolson (2020) offers insights into the connection between decision-making and feedback processing within continuous environments. The enhancement of the feedback-locked P300 following exploration, suggests its role in managing the explore-exploit trade-off.

The degree to which the cycle of decision-making generalizes from discrete to continuous tasks is not yet entirely clear. It is conceivable that the same cycle operates at various levels. Kolling and O'Reilly (2018) propose a differentiation between within-state and state-change decisions. Within-state decisions involve selecting the optimal course of action within a given environment, whereas state-change decisions entail comparing currently available options with potential alternatives upon a state transition. This framework indeed treats reward rate as a function of continuous time. Further

studies are needed to extend and better understand the cycle of decision making and feedback processing to more complex, everyday scenarios with sequential decision steps.

5.5 Future research

On a conceptual level, there remains uncertainty about the generalizability of MVT-type models to different types of foraging choices. Although the present work provides evidence by offering results on a foraging paradigm employing a continuous harvest phase, there is still a need for converging evidence across various types of tasks (Helversen et al., 2018). Moreover, a lack of ecological validity arises due to dissimilarities between these tasks and real-life foraging scenarios (Mobbs et al., 2018). Furthermore, decision neuroscience has yielded intriguing findings on inter-individual differences and the well-established phenomenon of risk aversion (Khaw et al., 2021). However, when it comes to foraging decisions, these topics have been largely neglected, leaving ample room for future studies to investigate in order to further the understanding of how individuals differ in their decision-making strategies and the role of risk aversion in foraging choices.

From a methodological point of view, the cluster-based permutation analyses used to correct for multiple comparisons in study 1 is a conservative method to extract significant regions in the time-frequency space. Moreover, cluster-based permutation analyses are dependent on the time window on which we focus. Because we focus on fairly large time windows within the foraging study, whereas the cluster-based permutation analyses being dependent on the time window, we cannot rule out the possibility of missing smaller effects and effects that fall outside the time window we examined. In addition, we used two different perspectives: stimulus-locked to represent foraging initiation phase and response-locked to represent the decision to leave. Since foraging decisions naturally take different lengths of time, it would be interesting in future projects to map the entire foraging process by either breaking down each foraging decision into individual reward steps or by using scaling techniques (Hassall et al., 2022) that stretch and compress the EEG signal to compare different trial lengths. When it comes to defining values related to a foraging decision, future studies should consider

the differentiation between the value of switching to a new patch versus the conflict experienced between choosing to stay or leave (Fontanesi et al., 2022). Simulations suggest that both processes provide distinct functions within anterior cingulate cortex (ACC; Brown & Alexander, 2017).

Apart from methodological considerations aimed at enhancing the understanding of decision-making and feedback processing, investigating the behavior of individuals with psychiatric disorders could provide valuable insights. Addicott et al. (2017) have highlighted the potential relevance of studying explore/exploit decision-making in such populations. Some psychiatric disorders may impact optimal foraging behavior. For example, patients with OCD tend to exhibit behavioral inflexibility, prioritize immediate rewards over long-term rewards (Cavedini et al., 2006), and require more information before making decisions (Foa et al., 2003). These traits suggest that individuals with OCD might display overexploitation of current options while neglecting long-term goals. Moreover, the assumed underlying neurological basis of OCD, the cortico-striato-thalamo-cortical circuit imbalance, which involves the posterior medial frontal cortex (pMFC), supports the hypothesis of the presence of impairments in foraging decisions. A study by Scholl et al. (2022) found that compulsivity was associated with searching for too long in the hope of getting a better offer due to insensitivity to costs, which lead to avoidance of situations where this bias could manifest. Further research is required to elucidate whether and how impairments in foraging choices manifest in patients with OCD. By studying how psychiatric disorders affect foraging decisions, we can gain a deeper understanding of decision-making processes in these populations and their potential neural correlates. By examining the neural correlates of feedback processing, which have already been extensively studied in psychiatric samples (Bellato et al., 2021; Endrass et al., 2013), we can effectively assess the efficacy of this approach. Impairments and biases in decision making can pose significant problems for those affected in their daily lives and make recovery difficult, so it is important to identify these, which consequently allows for the development of interventions.

Monitoring multiple reward rates at varying temporal scales (e.g., FRR and BRR), updating internal states through feedback, and adapting to a volatile environment pose a challenging and complex task. Current research, as noted by Grossman and Cohen

(2022), may not yet fully capture these dynamics, particularly those occurring on intermediate timescales. This also includes foraging decisions, which involve processes that occur "slowly" over seconds to minutes. Decision-making and learning are often continuous processes that operate within this middle timescale. Neuromodulators may play a crucial role in supporting and modifying activity within this timescale. Neurons can exhibit multiple timescales through synaptic and extrasynaptic neurotransmitter release, presenting an opportunity to tease apart the tracking of different time scales such as the FRR and BRR and to learn more about neuronal representation. Therefore, a topic worthy of expansion is the involvement of neuromodulators like dopamine in explore/exploit decisions, as has been addressed in some studies (Chakroun et al., 2020; Chakroun et al., 2022; Le Heron et al., 2020).

5.6 Conclusion

Previous research has demonstrated that animals efficiently navigate the exploration-exploitation dilemma while foraging for food in volatile environments, seeking to maximize rewards and minimize energy expenditure. Similarly, humans face such foraging decisions in real-life scenarios and learn from the outcomes. However, our understanding of the underlying mechanisms driving these processes and their neural correlates remains limited.

This thesis sheds light on how individuals explore and allocate resources in uncertain and dynamic environments, process feedback, and the corresponding neural correlates. Rather than limiting our perspective to simple decisions between distinct options, we must consider complex, sequential decisions that reflect real-world scenarios. The evidence presented here shows that environmental and current value representations manifest as distinct patterns of neuronal oscillations, tracking signals until a certain threshold is reached, aligning with normative foraging models. Moreover, neuronal feedback-related processes are influenced by the unexpectedness and valence of the feedback while representing various prediction errors.

By comprehending this decision-making cycle and adaptive behavior through feedback, we gain the ability to predict human choices and biases, enabling interventions to enhance decision-making and learning.

References

Acuña, D. E., & Schrater, P. (2010). Structure learning in human sequential decision-making. *PLoS Computational Biology*, *6*(12), e1001003. https://doi.org/10.1371/journal.pcbi.1001003

Addicott, M. A., Pearson, J. M., Sweitzer, M. M., Barack, D. L., & Platt, M. L. (2017). A Primer on Foraging and the Explore/Exploit Trade-Off for Psychiatry Research. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *42*(10), 1931–1939. https://doi.org/10.1038/npp.2017.108

Albrecht, C., van de Vijver, R., & Bellebaum, C. (2023). Learning new words via feedback-Association between feedback-locked ERPs and recall performance-An exploratory study. *Psychophysiology*, e14324. https://doi.org/10.1111/psyp.14324

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–1344. https://doi.org/10.1038/nn.2921

Arbel, Y., Goforth, K., & Donchin, E. (2013). The good, the bad, or the useful? The examination of the relationship between the feedback-related negativity (FRN) and long-term learning outcomes. *Journal of Cognitive Neuroscience*, *25*(8), 1249–1260. https://doi.org/10.1162/jocn_a_00385

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology*, *87*(1), 25–34. https://doi.org/10.1016/j.biopsycho.2011.01.010

Barack, D. L., & Gold, J. I. (2016). Temporal trade-offs in psychophysics. *Current Opinion in Neurobiology*, *37*, 121–125. https://doi.org/10.1016/j.conb.2016.01.015

Bellato, A., Norman, L., Idrees, I., Ogawa, C. Y., Waitt, A., Zuccolo, P. F., Tye, C., Radua, J., Groom, M. J., & Shephard, E. (2021). A systematic review and meta-analysis of altered electrophysiological markers of performance monitoring in Obsessive-Compulsive Disorder (OCD), Gilles de la Tourette Syndrome (GTS), Attention-Deficit/Hyperactivity disorder (ADHD) and Autism. *Neuroscience and Biobehavioral Reviews*, *131*, 964–987. https://doi.org/10.1016/j.neubiorev.2021.10.018

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *The European Journal of Neuroscience*, *27*(7), 1823–1835. https://doi.org/10.1111/j.1460-9568.2008.06138.x

Bellebaum, C., Polezzi, D., & Daum, I. (2010). It is less than you expected: The feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia*, *48*(11), 3343–3350. https://doi.org/10.1016/j.neuropsychologia.2010.07.023

Bernat, E. M., Nelson, L. D., & Baskin-Sommers, A. R. (2015). Time-frequency theta and delta measures index separable components of feedback processing in a gambling task. *Psychophysiology*, *52*(5), 626–637. https://doi.org/10.1111/psyp.12390

Bernat, E. M., Nelson, L. D., Holroyd, C. B., Gehring, W. J., & Patrick, C. J. (2008). Separating cognitive processes with principal components analysis of EEG time-frequency distributions. In F. T. Luk (Ed.), *SPIE Proceedings, Advanced Signal Processing Algorithms, Architectures, and Implementations XVIII* (70740S). SPIE. https://doi.org/10.1117/12.801362

Bernat, E. M., Nelson, L. D., Steele, V. R., Gehring, W. J., & Patrick, C. J. (2011). Externalizing psychopathology and gain-loss feedback in a simulated gambling task: Dissociable components of brain response revealed by time-frequency analysis. *Journal of Abnormal Psychology*, *120*(2), 352–364. https://doi.org/10.1037/a0022124

Brands, A. M., Mathar, D., & Peters, J. (2023). *Signatures of heuristic-based directed exploration in two-step sequential decision task behaviour*. preprint. https://doi.org/10.1101/2023.05.22.541443

Brosnan, M. B., Sabaroedin, K., Silk, T., Genc, S., Newman, D. P., Loughnane, G. M., Fornito, A., O'Connell, R. G., & Bellgrove, M. A. (2020). Evidence accumulation during perceptual decisions in humans varies as a function of dorsal frontoparietal organization. *Nature Human Behaviour*, *4*(8), 844–855. https://doi.org/10.1038/s41562-020-0863-4

Brown, J. W., & Alexander, W. H. (2017). Foraging Value, Risk Avoidance, and Multiple Control Signals: How the Anterior Cingulate Cortex Controls Value-based Decision-making. *Journal of Cognitive Neuroscience*, *29*(10), 1656–1673. https://doi.org/10.1162/jocn_a_01140

Bruijn, E. R. A. de, Mars, R. B., & Hulstijn, W. (2004). 'It wasn't me… or was it?' How false feedback effects performance. In M. Ullsperger & M. Falkenstein (Eds.), *MPI special issue in human cognitive and brain sciences. Errors, conflicts, and the brain. Current opinions on performance monitoring* (pp. 118–124). Max Planck Institute of Cognitive Neuroscience. https://hdl.handle.net/2066/64750

Burnside, R., Fischer, A. G., & Ullsperger, M. (2019). The feedback-related negativity indexes prediction error in active but not observational learning. *Psychophysiology*, *56*(9), e13389. https://doi.org/10.1111/psyp.13389

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and Neural Bases of Multi-Attribute, Multi-Alternative, Value-based Decisions. *Trends in Cognitive Sciences*, *23*(3), 251–263. https://doi.org/10.1016/j.tics.2018.12.003

Cavanagh, J. F. (2015). Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *NeuroImage*, *110*, 205–216. https://doi.org/10.1016/j.neuroimage.2015.02.007

Cavanagh, J. F., Bismark, A. W., Frank, M. J., & Allen, J. J. B. (2019). Multiple Dissociations Between Comorbid Depression and Anxiety on Reward and Punishment Processing: Evidence From Computationally Informed EEG.

Computational Psychiatry (Cambridge, Mass.), *3*, 1–17. https://doi.org/10.1162/cpsy_a_00024

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex (New York, N.Y. : 1991)*, *22*(11), 2575–2586. https://doi.org/10.1093/cercor/bhr332

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, *49*(4), 3198–3209. https://doi.org/10.1016/j.neuroimage.2009.11.080

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence. *Journal of Physiology, Paris*, *109*(1-3), 3–15. https://doi.org/10.1016/j.jphysparis.2014.04.003

Cavedini, P., Gorini, A., & Bellodi, L. (2006). Understanding obsessive-compulsive disorder: Focus on decision making. *Neuropsychology Review*, *16*(1), 3–15. https://doi.org/10.1007/s11065-006-9001-y

Chakroun, K., Mathar, D., Wiehler, A., Ganzer, F., & Peters, J. (2020). Dopaminergic modulation of the exploration/exploitation trade-off in human decision-making. *ELife*, *9*. https://doi.org/10.7554/eLife.51260

Chakroun, K., Wiehler, A., Wagner, B., Mathar, D., Ganzer, F., vanEimeren, T., Sommer, T., & Peters, J. (2022). *Dopamine regulates decision thresholds in human reinforcement learning*. preprint. https://doi.org/10.1101/2022.09.29.509499

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. *Journal of Cognitive Neuroscience*, *23*(4), 936–946. https://doi.org/10.1162/jocn.2010.21456

Clairis, N., & Lopez-Persem, A. (2023). Debates on the dorsomedial prefrontal/dorsal anterior cingulate cortex: Insights for future research. *Brain : A Journal of Neurology*. Advance online publication. https://doi.org/10.1093/brain/awad263

Cockburn, J., & Holroyd, C. B. (2018). Feedback information and the reward positivity. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *132*(Pt B), 243–251. https://doi.org/10.1016/j.ijpsycho.2017.11.017

Cogliati Dezza, I., Yu, A. J., Cleeremans, A., & Alexander, W. H. (2017). Learning the value of information and reward over time when solving exploration-exploitation problems. *Scientific Reports*, *7*(1), 16919. https://doi.org/10.1038/s41598-017-17237-w

Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *The Journal of Neuroscience : The Official Journal of the Society for*

Neuroscience, *27*(2), 371–378. https://doi.org/10.1523/JNEUROSCI.4421-06.2007

Collins, A. G. E., & Cockburn, J. (2020). Beyond dichotomies in reinforcement learning. *Nature Reviews. Neuroscience*, *21*(10), 576–586. https://doi.org/10.1038/s41583-020-0355-6

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective & Behavioral Neuroscience*, *15*(4), 837–853. https://doi.org/10.3758/s13415-015-0350-y

Cortes, P. M., García-Hernández, J. P., Iribe-Burgos, F. A., Hernández-González, M., Sotelo-Tapia, C., & Guevara, M. A. (2021). Temporal division of the decision-making process: An EEG study. *Brain Research*, *1769*, 147592. https://doi.org/10.1016/j.brainres.2021.147592

Courchesne, E., Hillyard, S. A., & Courchesne, R. Y. (1977). P3 waves to the discrimination of targets in homogeneous and heterogeneous stimulus sequences. *Psychophysiology*, *14*(6), 590–597. https://doi.org/10.1111/j.1469-8986.1977.tb01206.x

Danielmeier, C., Allen, E. A., Jocham, G., Onur, O. A., Eichele, T., & Ullsperger, M. (2015). Acetylcholine mediates behavioral and neural post-error control. *Current Biology : CB*, *25*(11), 1461–1468. https://doi.org/10.1016/j.cub.2015.04.022

Danielmeier, C., Eichele, T., Forstmann, B. U., Tittgemeyer, M., & Ullsperger, M. (2011). Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(5), 1780–1789. https://doi.org/10.1523/JNEUROSCI.4299-10.2011

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, *2*, 233. https://doi.org/10.3389/fpsyg.2011.00233

Davidson, J. D., & El Hady, A. (2019). Foraging as an evidence accumulation process. *PLoS Computational Biology*, *15*(7), e1007060. https://doi.org/10.1371/journal.pcbi.1007060

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Cramon, D. Y. von, & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *25*(50), 11730–11737. https://doi.org/10.1523/JNEUROSCI.3286-05.2005

Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, *34*(4), 1443–1449. https://doi.org/10.1016/j.neuroimage.2006.11.004.

Dennison, J. B., Sazhin, D., & Smith, D. V. (2022). Decision neuroscience and neuroeconomics: Recent progress and ongoing challenges. *Wiley Interdisciplinary Reviews. Cognitive Science*, *13*(3), e1589. https://doi.org/10.1002/wcs.1589

Domenech, P., Rheims, S., & Koechlin, E. (2020). Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. *Science (New York, N.Y.)*, *369*(6507). https://doi.org/10.1126/science.abb0184

Donchin, E., & Coles, M. G. H. (1998). Context updating and the P300. *Behavioral and Brain Sciences*, *21*(1), 152–154. https://doi.org/10.1017/S0140525X98230950

Donkers, F. C., & van Boxtel, G. J. (2005). Mediofrontal Negativities to Averted Gains and Losses in the Slot-Machine Task. *Journal of Psychophysiology*, *19*(4), 256–262. https://doi.org/10.1027/0269-8803.19.4.256

Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Current Biology : CB*, *19*(18), 1581–1585. https://doi.org/10.1016/j.cub.2009.07.066

Donoso, M., Collins, A. G. E., & Koechlin, E. (2014). Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science (New York, N.Y.)*, *344*(6191), 1481–1486. https://doi.org/10.1126/science.1252254

Endrass, T., Koehne, S., Riesel, A., & Kathmann, N. (2013). Neural correlates of feedback processing in obsessive-compulsive disorder. *Journal of Abnormal Psychology*, *122*(2), 387–396. https://doi.org/10.1037/a0031496

Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(35), 12087–12092. https://doi.org/10.1523/JNEUROSCI.1410-12.2012

Findling, C., Chopin, N., & Koechlin, E. (2021). Imprecise neural computations as a source of adaptive behaviour in volatile environments. *Nature Human Behaviour*, *5*(1), 99–112. https://doi.org/10.1038/s41562-020-00971-z

Fischer, A. G., Danielmeier, C., Villringer, A., Klein, T. A., & Ullsperger, M. (2016). Gender Influences on Brain Responses to Errors and Post-Error Adjustments. *Scientific Reports*, *6*, 24435. https://doi.org/10.1038/srep24435

Fischer, A. G., Nigbur, R., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2018). Cortical beta power reflects decision dynamics and uncovers multiple facets of post-error adaptation. *Nature Communications*, *9*(1), 5038. https://doi.org/10.1038/s41467-018-07456-8

Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006

Foa, E. B., Mathews, A., Abramowitz, J. S., Amir, N., Przeworski, A., Riggs, D. S., Filip, J. C., & Alley, A. (2003). Do patients with obsessive–compulsive disorder have deficits in decision-making? *Cognitive Therapy and Research*, *27*(4), 431–445. https://doi.org/10.1023/A:1025424530644

Fontanesi, L., Shenhav, A., & Gluth, S. (2022). Disentangling choice value and choice conflict in sequential decisions under risk. *PLoS Computational Biology*, *18*(10), e1010478. https://doi.org/10.1371/journal.pcbi.1010478

Foti, D., & Weinberg, A. (2018). Reward and feedback processing: State of the field, best practices, and future directions. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *132*(Pt B), 171–174. https://doi.org/10.1016/j.ijpsycho.2018.08.006

Foti, D., Weinberg, A., Bernat, E. M., & Proudfit, G. H. (2015). Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *126*(7), 1338–1347. https://doi.org/10.1016/j.clinph.2014.08.025

Frömer, R., Dean Wolf, C. K., & Shenhav, A. (2019). Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making. *Nature Communications*, *10*(1), 4926. https://doi.org/10.1038/s41467-019-12931-x

Gabay, A. S., & Apps, M. A. J. (2021). Foraging optimally in social neuroscience: Computations and methodological considerations. *Social Cognitive and Affective Neuroscience*, *16*(8), 782–794. https://doi.org/10.1093/scan/nsaa037

Garrett, N., & Daw, N. D. (2020). Biased belief updating and suboptimal choice in foraging decisions. *Nature Communications*, *11*(1), 3417. https://doi.org/10.1038/s41467-020-16964-5

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science (New York, N.Y.)*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement. *NeuroImage*, *47*(4), 2023–2030. https://doi.org/10.1016/j.neuroimage.2009.05.064

Gheza, D., Paul, K., & Pourtois, G. (2018). Dissociable effects of reward and expectancy during evaluative feedback processing revealed by topographic ERP mapping analysis. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *132*(Pt B), 213–225. https://doi.org/10.1016/j.ijpsycho.2017.11.013

Glazer, J., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *132*(Pt B), 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Glazer, J., & Nusslock, R. (2021). Outcome valence and stimulus frequency affect neural responses to rewards and punishments. *Psychophysiology*, e13981. https://doi.org/10.1111/psyp.13981

Gluth, S., Rieskamp, J., & Büchel, C. (2014). Neural evidence for adaptive strategy selection in value-based decision-making. *Cerebral Cortex (New York, N.Y. : 1991)*, *24*(8), 2009–2021. https://doi.org/10.1093/cercor/bht049

Grossman, C. D., & Cohen, J. Y. (2022). Neuromodulation and Neurophysiology on the Timescale of Learning and Decision-Making. *Annual Review of Neuroscience*, *45*, 317–337. https://doi.org/10.1146/annurev-neuro-092021-125059

Gruendler, T. O. J., Ullsperger, M., & Huster, R. J. (2011). Event-related potential correlates of performance-monitoring in a lateralized time-estimation task. *PloS One*, *6*(10), e25591. https://doi.org/10.1371/journal.pone.0025591

Guan, Q., Ma, L., Chen, Y., Luo, Y., & He, H. (2023). Midfrontal theta phase underlies evidence accumulation and response thresholding in cognitive control. *Cerebral Cortex (New York, N.Y. : 1991)*. Advance online publication. https://doi.org/10.1093/cercor/bhad175

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*(2), 161–170. https://doi.org/10.1111/j.1469-8986.2005.00278.x

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Harhen, N. C., & Bornstein, A. M. (2023). Overharvesting in human patch foraging reflects rational structure learning and adaptive planning. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(13), e2216524120. https://doi.org/10.1073/pnas.2216524120

Hassall, C. D., Harley, J., Kolling, N., & Hunt, L. T. (2022). Temporal scaling of human scalp-recorded potentials. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(43), e2214638119. https://doi.org/10.1073/pnas.2214638119

Hassall, C. D., & Krigolson, O. E. (2020). Feedback processing is enhanced following exploration in continuous environments. *Neuropsychologia*, *146*, 107538. https://doi.org/10.1016/j.neuropsychologia.2020.107538

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), 933–939. https://doi.org/10.1038/nn.2856

Helversen, B. von, Mata, R., Samanez-Larkin, G. R., & Wilke, A. (2018). Foraging, exploration, or search? On the (lack of) convergent validity between three behavioral paradigms. *Evolutionary Behavioral Sciences*, *12*(3), 152–162. https://doi.org/10.1037/ebs0000121

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037//0033-295X.109.4.679

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917. https://doi.org/10.1111/j.1469-8986.2007.00561.x

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Hoy, C. W., Steiner, S. C., & Knight, R. T. (2021). Single-trial modeling separates multiple overlapping prediction errors during reward processing in human EEG. *Communications Biology*, *4*(1), 910. https://doi.org/10.1038/s42003-021-02426-1

Hunt, L. T., Daw, N. D., Kaanders, P., MacIver, M. A., Mugan, U., Procyk, E., Redish, A. D., Russo, E., Scholl, J., Stachenfeld, K., Wilson, C. R. E., & Kolling, N. (2021). Formalizing planning and information search in naturalistic decision-making. *Nature Neuroscience*, *24*(8), 1051–1064. https://doi.org/10.1038/s41593-021-00866-w

Hutchinson, J. M., Wilke, A., & Todd, P. M. (2008). Patch leaving in humans: can a generalist adapt its rules to dispersal of items across patches? *Animal Behaviour*, *75*(4), 1331–1349. https://doi.org/10.1016/j.anbehav.2007.09.006

Huvermann, D. M., Bellebaum, C., & Peterburs, J. (2021). Selective Devaluation Affects the Processing of Preferred Rewards. *Cognitive, Affective & Behavioral Neuroscience*, *21*(5), 1010–1025. https://doi.org/10.3758/s13415-021-00904-x

Intriligator, J., & Polich, J. (1994). On the relationship between background EEG and the P300 event-related potential. *Biological Psychology*, *37*(3), 207–218. https://doi.org/10.1016/0301-0511(94)90003-5

Jacobs, J., Hwang, G., Curran, T., & Kahana, M. J. (2006). Eeg oscillations and recognition memory: Theta correlates of memory retrieval and decision making. *NeuroImage*, *32*(2), 978–987. https://doi.org/10.1016/j.neuroimage.2006.02.018

Jepma, M., Brown, S. B. R. E., Murphy, P. R., Koelewijn, S. C., Vries, B. de, van den Maagdenberg, A. M., & Nieuwenhuis, S. (2018). Noradrenergic and Cholinergic Modulation of Belief Updating. *Journal of Cognitive Neuroscience*, *30*(12), 1803–1820. https://doi.org/10.1162/jocn_a_01317

Jepma, M., Murphy, P. R., Nassar, M. R., Rangel-Gomez, M., Meeter, M., & Nieuwenhuis, S. (2016). Catecholaminergic Regulation of Learning Rate in a Dynamic Environment. *PLoS Computational Biology*, *12*(10), e1005171. https://doi.org/10.1371/journal.pcbi.1005171

Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(5), 1606–1613. https://doi.org/10.1523/JNEUROSCI.3904-10.2011

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. *The*

Journal of Neuroscience : The Official Journal of the Society for Neuroscience, *29*(23), 7489–7496. https://doi.org/10.1523/JNEUROSCI.0349-09.2009

Johnson, R. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, *23*(4), 367–384. https://doi.org/10.1111/j.1469-8986.1986.tb00649.x

Johnson, R., & Donchin, E. (1980). P300 and stimulus categorization: Two plus one is not so different from one plus one. *Psychophysiology*, *17*(2), 167–178. https://doi.org/10.1111/j.1469-8986.1980.tb00131.x

Kaiser, L. F., Gruendler, T. O. J., Speck, O., Luettgau, L., & Jocham, G. (2021). Dissociable roles of cortical excitation-inhibition balance during patch-leaving versus value-guided decisions. *Nature Communications*, *12*(1), 904. https://doi.org/10.1038/s41467-020-20875-w

Kamarajan, C., Porjesz, B., Rangaswamy, M., Tang, Y., Chorlian, D. B., Padmanabhapillai, A., Saunders, R., Pandey, A. K., Roopesh, B. N., Manz, N., Stimus, A. T., & Begleiter, H. (2009). Brain signatures of monetary loss and gain: Outcome-related potentials in a single outcome gambling task. *Behavioural Brain Research*, *197*(1), 62–76. https://doi.org/10.1016/j.bbr.2008.08.011

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*(7), 940–947. https://doi.org/10.1038/nn1724

Keren, H., O'Callaghan, G., Vidal-Ribas, P., Buzzell, G. A., Brotman, M. A., Leibenluft, E., Pan, P. M., Meffert, L., Kaiser, A., Wolke, S., Pine, D. S., & Stringaris, A. (2018). Reward Processing in Depression: A Conceptual and Meta-Analytic Review Across fMRI and EEG Studies. *The American Journal of Psychiatry*, *175*(11), 1111–1120. https://doi.org/10.1176/appi.ajp.2018.17101124

Khanna, P., & Carmena, J. M. (2015). Neural oscillations: Beta band activity across motor networks. *Current Opinion in Neurobiology*, *32*, 60–67. https://doi.org/10.1016/j.conb.2014.11.010

Khaw, M. W., Li, Z., & Woodford, M. (2021). Cognitive Imprecision and Small-Stakes Risk Aversion. *The Review of Economic Studies*, *88*(4), 1979–2013. https://doi.org/10.1093/restud/rdaa044

Kilpatrick, Z. P., Davidson, J. D., & El Hady, A. (2021). Uncertainty drives deviations in normative foraging decision strategies. *Journal of the Royal Society, Interface*, *18*(180), 20210337. https://doi.org/10.1098/rsif.2021.0337

King, J. A., Korb, F. M., Cramon, D. Y. von, & Ullsperger, M. (2010). Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(38), 12759–12769. https://doi.org/10.1523/JNEUROSCI.3274-10.2010

Kirschner, H., Fischer, A. G., Danielmeier, C., Klein, T. A., & Ullsperger, M. (2023). *Cortical beta power reflects a neural implementation of decision boundary collapse in speeded decisions*. preprint. https://doi.org/10.1101/2023.01.13.523918

Kirschner, H., Fischer, A. G., & Ullsperger, M. (2022). Feedback-related EEG dynamics separately reflect decision parameters, biases, and future choices. *NeuroImage*, *259*, 119437. https://doi.org/10.1016/j.neuroimage.2022.119437

Kolling, N., & Akam, T. (2017). (Reinforcement?) Learning to forage optimally. *Current Opinion in Neurobiology*, *46*, 162–169. https://doi.org/10.1016/j.conb.2017.08.008

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural mechanisms of foraging. *Science (New York, N.Y.)*, *336*(6077), 95–98. https://doi.org/10.1126/science.1216930

Kolling, N., & O'Reilly, J. X. (2018). State-change decisions and dorsomedial prefrontal cortex: The importance of time. *Current Opinion in Behavioral Sciences*, *22*, 152–160. https://doi.org/10.1016/j.cobeha.2018.06.017

Kolling, N., Wittmann, M. K., Behrens, T. E. J., Boorman, E. D., Mars, R. B., & Rushworth, M. F. S. (2016). Value, search, persistence and model updating in anterior cingulate cortex. *Nature Neuroscience*, *19*(10), 1280–1285. https://doi.org/10.1038/nn.4382

Krigolson, O. E. (2018). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *132*(Pt B), 175–183. https://doi.org/10.1016/j.ijpsycho.2017.11.007

Kristjánsson, Á., Ólafsdóttir, I. M., & Kristjánsson, T. (2020). Visual Foraging Tasks Provide New Insights into the Orienting of Visual Attention: Methodological Considerations. In S. Pollmann (Ed.), *Neuromethods. Spatial Learning and Attention Guidance* (Vol. 151, pp. 3–21). Springer US. https://doi.org/10.1007/7657_2019_21

Laureiro-Martínez, D., Brusoni, S., Canessa, N., & Zollo, M. (2015). Understanding the exploration-exploitation dilemma: An fMRI study of attention control and decision-making performance. *Strategic Management Journal*, *36*(3), 319–338. https://doi.org/10.1002/smj.2221

Le Heron, C., Kolling, N., Plant, O., Kienast, A., Janska, R., Ang, Y.-S., Fallon, S., Husain, M., & Apps, M. A. J. (2020). Dopamine Modulates Dynamic Decision-Making during Foraging. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *40*(27), 5273–5282. https://doi.org/10.1523/JNEUROSCI.2586-19.2020

Lee, S., Yu, L. Q., Lerman, C., & Kable, J. W. (2021). Subjective value, not a gridlike code, describes neural activity in ventromedial prefrontal cortex during value-based decision-making. *NeuroImage*, *237*, 118159. https://doi.org/10.1016/j.neuroimage.2021.118159

Liu, T., & Pleskac, T. J. (2011). Neural correlates of evidence accumulation in a perceptual decision task. *Journal of Neurophysiology*, *106*(5), 2383–2398. https://doi.org/10.1152/jn.00413.2011

Liu, Y., Nelson, L. D., Bernat, E. M., & Gehring, W. J. (2014). Perceptual properties of feedback stimuli influence the feedback-related negativity in the flanker gambling task. *Psychophysiology*, *51*(8), 782–788. https://doi.org/10.1111/psyp.12216

Mansouri, F. A., Koechlin, E., Rosa, M. G. P., & Buckley, M. J. (2017). Managing competing goals - a key role for the frontopolar cortex. *Nature Reviews. Neuroscience*, *18*(11), 645–657. https://doi.org/10.1038/nrn.2017.111

Marco-Pallarés, J., Camara, E., Münte, T. F., & Rodríguez-Fornells, A. (2008). Neural mechanisms underlying adaptive actions after slips. *Journal of Cognitive Neuroscience*, *20*(9), 1595–1610. https://doi.org/10.1162/jocn.2008.20117

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(47), 12539–12545. https://doi.org/10.1523/JNEUROSCI.2925-08.2008

McGuire, J. T., & Kable, J. W. (2015). Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nature Neuroscience*, *18*(5), 760–766. https://doi.org/10.1038/nn.3994

Miltner, W. H., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Mobbs, D., Trimmer, P. C., Blumstein, D. T., & Dayan, P. (2018). Foraging for foundations in decision neuroscience: Insights from ethology. *Nature Reviews. Neuroscience*, *19*(7), 419–427. https://doi.org/10.1038/s41583-018-0010-7

Nácher, V., Ledberg, A., Deco, G., & Romo, R. (2013). Coherent delta-band oscillations between cortical areas correlate with decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(37), 15085–15090. https://doi.org/10.1073/pnas.1314681110

Nassar, M. R., Bruckner, R., & Frank, M. J. (2019). Statistical context dictates the relationship between feedback-related EEG signals and learning. *ELife*, *8*. https://doi.org/10.7554/eLife.46975

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25. https://doi.org/10.1002/hbm.1058

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, *131*(4), 510–532. https://doi.org/10.1037/0033-2909.131.4.510

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neuroscience and Biobehavioral Reviews*, *28*(4), 441–448. https://doi.org/10.1016/j.neubiorev.2004.05.003

O'Connell, R. G., & Kelly, S. P. (2021). Neurophysiology of Human Perceptual Decision-Making. *Annual Review of Neuroscience*, *44*, 495–516. https://doi.org/10.1146/annurev-neuro-092019-100200

Odum, A. L. (2011). Delay Discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior*, *96*(3), 427–439. https://doi.org/10.1901/jeab.2011.96-423

Overmeyer, R., Kirschner, H., Fischer, A. G., & Endrass, T. (2023). *Motivation Matters: Unraveling the Influence of Trial-Based Motivational Changes on Performance Monitoring Stages in a Flanker Task*. preprint. https://doi.org/10.21203/rs.3.rs-3087426/v1

Palmer, J., Kreutz-Delgado, K., & Makeig, S. (2012). AMICA: An adaptive mixture of 13 independent component analyzers with shared components. *Swartz Center for Computatonal Neursoscience, University of California San Diego, Tech. Rep.*

Pape, A.-A., & Siegel, M. (2016). Motor cortex activity predicts response alternation during sensorimotor decisions. *Nature Communications*, *7*, 13098. https://doi.org/10.1038/ncomms13098

Pereira, M., Megevand, P., Tan, M. X., Chang, W., Wang, S., Rezai, A., Seeck, M., Corniola, M., Momjian, S., Bernasconi, F., Blanke, O., & Faivre, N. (2021). Evidence accumulation relates to perceptual consciousness and monitoring. *Nature Communications*, *12*(1), 3261. https://doi.org/10.1038/s41467-021-23540-y

Peters, J., & Büchel, C. (2009). Overlapping and Distinct Neural Systems Code for Subjective Value during Intertemporal and Risky Decision Making. *The Journal of Neuroscience*, *29*(50), 15727–15734. https://doi.org/10.1523/JNEUROSCI.3489-09.2009

Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(44), 11912–11924. https://doi.org/10.1523/JNEUROSCI.3522-07.2007

Polanía, R., Krajbich, I., Grueschow, M., & Ruff, C. C. (2014). Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron*, *82*(3), 709–720. https://doi.org/10.1016/j.neuron.2014.03.014

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, *52*(4), 449–459. https://doi.org/10.1111/psyp.12370

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews. Neuroscience*, *9*(7), 545–556. https://doi.org/10.1038/nrn2357

Rogge, J., Jocham, G., & Ullsperger, M. (2022). Motor cortical signals reflecting decision making and action preparation. *NeuroImage*, *263*, 119667. https://doi.org/10.1016/j.neuroimage.2022.119667

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. https://doi.org/10.1037/bul0000006

Sambrook, T. D., & Goslin, J. (2016). Principal components analysis of reward prediction errors in a reinforcement learning task. *NeuroImage*, *124*(Pt A), 276–286. https://doi.org/10.1016/j.neuroimage.2015.07.032

San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, *6*, 304. https://doi.org/10.3389/fnhum.2012.00304

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport*, *16*(4), 407–411. https://doi.org/10.1097/00001756-200503150-00020

Scholl, J., Trier, H. A., Rushworth, M. F. S., & Kolling, N. (2022). The effect of apathy and compulsivity on planning and stopping in sequential decision-making. *PLoS Biology*, *20*(3), e3001566. https://doi.org/10.1371/journal.pbio.3001566

Severo, M. C., Walentowska, W., Moors, A., & Pourtois, G. (2018). Goals matter: Amplification of the motivational significance of the feedback when goal impact is increased. *Brain and Cognition*, *128*, 56–72. https://doi.org/10.1016/j.bandc.2018.11.002

Shadmehr, R., & Ahmed, A. A. (2020). *Vigor: Neuroeconomics of movement control*. The MIT Press. https://doi.org/10.7551/mitpress/12940.001.0001?locatt=mode:legacy

Sharot, T., Guitart-Masip, M., Korn, C. W., Chowdhury, R., & Dolan, R. J. (2012). How dopamine enhances an optimism bias in humans. *Current Biology : CB*, *22*(16), 1477–1481. https://doi.org/10.1016/j.cub.2012.05.053

Shenhav, A., Straccia, M. A., Musslick, S., Cohen, J. D., & Botvinick, M. M. (2018). Dissociable neural mechanisms track evidence accumulation for selection of attention versus action. *Nature Communications*, *9*(1), 2485. https://doi.org/10.1038/s41467-018-04841-1

Sims, C. R., Neth, H., Jacobs, R. A., & Gray, W. D. (2013). Melioration as rational choice: Sequential decision making in uncertain environments. *Psychological Review*, *120*(1), 139–154. https://doi.org/10.1037/a0030850

Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science (New York, N.Y.)*, *193*(4258), 1142–1146. https://doi.org/10.1126/science.959831

Steixner-Kumar, S., & Gläscher, J. (2020). Strategies for navigating a dynamic world. *Science (New York, N.Y.)*, *369*(6507), 1056–1057. https://doi.org/10.1126/science.abd7258

Stewardson, H. J., & Sambrook, T. D. (2021). Reward, Salience, and Agency in Event-Related Potentials for Appetitive and Aversive Contexts. *Cerebral Cortex (New York, N.Y. : 1991)*, *31*(11), 5006–5014. https://doi.org/10.1093/cercor/bhab137/6279867

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(19), 8264–8269. https://doi.org/10.1523/JNEUROSCI.5695-12.2013

Toyomaki, A., & Murohashi, H. (2005). Discrepancy between feedback negativity and subjective evaluation in gambling. *Neuroreport*, *16*(16), 1865–1868. https://doi.org/10.1097/01.wnr.0000185962.96217.36

Ullsperger, M. (2017). Neural Bases of Performance Monitoring. In T. Egner (Ed.), *The Wiley Handbook of Cognitive Control* (pp. 292–313). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118920497.ch17

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences*, *18*(5), 259–267. https://doi.org/10.1016/j.tics.2014.02.009

van Boxtel, G. J. (2004). The use of the subtraction technique in the psychophysiology of response inhibition and conflict. In M. Ullsperger & M. Falkenstein (Eds.), *MPI special issue in human cognitive and brain sciences. Errors, conflicts, and the brain. Current opinions on performance monitoring* (pp. 219–225). Max Planck Institute of Cognitive Neuroscience. https://research.tilburguniversity.edu/en/publications/the-use-of-the-subtraction-technique-in-the-psychophysiology-of-r

van Noordt, S. J. R., & Segalowitz, S. J. (2012). Performance monitoring and the medial prefrontal cortex: A review of individual differences and context effects as a window on self-regulation. *Frontiers in Human Neuroscience*, *6*, 197. https://doi.org/10.3389/fnhum.2012.00197

Verharen, J. P. H., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2019). Differential contributions of striatal dopamine D1 and D2 receptors to component processes of value-based decision making. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *44*(13), 2195–2204. https://doi.org/10.1038/s41386-019-0454-0

Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology*, *34*(2), 131–156. https://doi.org/10.1111/j.1469-8986.1997.tb02125.x

Verleger, R., Jaskowski, P., & Wauschkuhn, B. (1994). Suspense and surprise: On the relationship between expectancies and P3. *Psychophysiology*, *31*(4), 359–369. https://doi.org/10.1111/j.1469-8986.1994.tb02444.x

Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., & Debener, S. (2009). Semi-automatic identification of independent components representing EEG artifact. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *120*(5), 868–877. https://doi.org/10.1016/j.clinph.2009.01.015

Walentowska, W., Moors, A., Paul, K., & Pourtois, G. (2016). Goal relevance influences performance monitoring at the level of the FRN and P3 components. *Psychophysiology*, *53*(7), 1020–1033. https://doi.org/10.1111/psyp.12651

Walentowska, W., Severo, M. C., Moors, A., & Pourtois, G. (2019). When the outcome is different than expected: Subjective expectancy shapes reward prediction error at the FRN level. *Psychophysiology*, *56*(12), e13456. https://doi.org/10.1111/psyp.13456

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Webb, C. A., Auerbach, R. P., Bondy, E., Stanton, C. H., Foti, D., & Pizzagalli, D. A. (2017). Abnormal neural responses to feedback in depressed adolescents. *Journal of Abnormal Psychology*, *126*(1), 19–31. https://doi.org/10.1037/abn0000228

Weismüller, B., & Bellebaum, C. (2016). Expectancy affects the feedback-related negativity (FRN) for delayed feedback in probabilistic learning. *Psychophysiology*, *53*(11), 1739–1750. https://doi.org/10.1111/psyp.12738

Williams, C. C., Ferguson, T. D., Hassall, C. D., Abimbola, W., & Krigolson, O. E. (2021). The ERP, frequency, and time-frequency correlates of feedback processing: Insights from a large sample study. *Psychophysiology*, *58*(2), e13722. https://doi.org/10.1111/psyp.13722

Wittmann, M. K., Kolling, N., Akaishi, R., Chau, B. K. H., Brown, J. W., Nelissen, N., & Rushworth, M. F. S. (2016). Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nature Communications*, *7*, 12327. https://doi.org/10.1038/ncomms12327

Wolfe, J. M. (2013). When is it time to move to the next raspberry bush? Foraging rules in human visual search. *Journal of Vision*, *13*(3), 10. https://doi.org/10.1167/13.3.10

Yan, Y., Hunt, L. T., & Hassall, C. D. (2023). *Reward positivity biases interval production in a continuous timing task*. preprint. https://doi.org/10.1101/2023.07.06.548049

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). Erp correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex (New York, N.Y. : 1991)*, *15*(5), 535–544. https://doi.org/10.1093/cercor/bhh153

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *24*(28), 6258–6264. https://doi.org/10.1523/JNEUROSCI.4537-03.2004

Zheng, Y., & Mei, S. (2023). Neural dissociation between reward and salience prediction errors through the lens of optimistic bias. *Human Brain Mapping*. Advance online publication. https://doi.org/10.1002/hbm.26398

Attachments

A Autor contributions

B Declaration of Honor

A Author contributions

Kirsch, F.[1], Kirschner, H.[1], Fischer, A. G., Klein, T. A.[2], & Ullsperger, M.[2] (2022). Disentangling performance-monitoring signals encoded in feedback-related EEG dynamics. *NeuroImage*, *257*, 119322. https://doi.org/10.1016/j.neuroimage.2022.119322

1 These authors contributed equally to this work.

2 TK and MU should be considered joint senior author.

Franziska Kirsch: Conceptualization, Methodology, Data curation, Formal analysis, Visualization, Writing –original draft. Hans Kirschner: Conceptualization, Methodology, Data curation, Formal analysis, Visualization, Writing –review & editing. Adrian G. Fischer: Software, Formal analysis, Methodology, Writing –review & editing. Tilmann A. Klein: Conceptualization, Methodology, Funding acquisition, Resources, Supervision, Writing –review & editing. Markus Ullsperger: Conceptualization, Methodology, Funding acquisition, Re- sources, Supervision, Writing –review & editing.

B Declaration of Honor

I hereby declare that I prepared this thesis without impermissible help of third parties and that none other than the indicated tools have been used; all sources of information are clearly marked, including my own publications. In particular I have not consciously:

- Fabricated data or rejected undesired results
- Misused statistical methods with the aim of drawing other conclusions than those warranted by the available data
- Plagiarized external data or publications
- Presented the results of other researchers in a distorted way

I am aware that violations of copyright may lead to injunction and damage claims of the author and also to prosecution by the law enforcement authorities. I hereby agree that the thesis may be reviewed for plagiarism by mean of electronic data processing. This work has not yet been submitted as a doctoral thesis in the same or a similar form in Germany or in any other country. It has not yet been published as a whole.

Leipzig, 29 August 2023

Franziska Kirsch