Received: 20 September 2019 | Revised: 31 January 2020 | Accepted: 5 February 2020
DOI: 10.1111/psyp.13557

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

It's costly punishment, not altruistic: Low midfrontal theta and state anger predict punishment

Johannes Rodrigues[1] | **Marvin Liesner**[1] | **Mario Reutter**[1] | **Patrick Mussel**[2] | **Johannes Hewig**[1]

[1]Institute of Psychology, Julius-Maximilians-Universität Würzburg, Würzburg, Germany
[2]Division Personality Psychology and Psychological Assessment, Freie Universität Berlin, Berlin, Germany

Correspondence
Johannes Rodrigues, Psychology I: Differential Psychology, Personality Psychology and Psychological Diagnostics, Julius-Maximilians-Universität Würzburg, Marcusstraße 9-11, Würzburg 97070, Germany.
Email: johannes.rodrigues@uni-wuerzburg.de

Funding information
This publication was funded by the European Union through the project "Individualisierung Digital" (Fonds 823881) in the Europäischer Fonds für regionale Entwicklung (EFRE)

Abstract
Punishment in economic games has been interpreted as "altruistic." However, it was shown that punishment is related to trait anger instead of trait altruism in a third-party dictator game if compensation is also available. Here, we investigated the influence of state anger on punishment and compensation in the third-party dictator game. Therefore, we used movie sequences for emotional priming, including the target states anger, happy, and neutral. We measured the Feedback-Related Negativity (FRN) and midfrontal theta band activation, to investigate an electro-cortical correlate of the processing of fair and unfair offers. Also, we assessed single-trial FRN and midfrontal theta band activation as a predictor for punishment and compensation. We found that punishment was linked to state anger. Midfrontal theta band activation, which has previously been linked to altruistic acts and cognitive control, predicted less punishment. Additionally, trait anger led to enhanced FRN for unfair offers. This led to the interpretation that the FRN depicts the evaluation of fairness, while midfrontal theta band activation captures an aspect of cognitive control and altruistic motivation. We conclude that we need to redefine "altruistic punishment" into "costly punishment," as no direct link of altruism and punishment is given. Additionally, midfrontal theta band activation complements the FRN and offers additional insights into complex responses and decision processes, especially as a single trial predictor.

KEYWORDS
altruistic compensation, costly punishment, FRN, midfrontal theta activation, third party dictator game

1 | INTRODUCTION

Altruistic behavior, the voluntary action intended to benefit another person without the expectancy of receiving external rewards or avoiding external punishments (Eisenberg & Miller, 1987), can have many motives. In order to give a more specific description of prosocial behavior types, Carlo and Randall (2002) used a factor analytical approach that also led to a more narrow definition of altruism (see also Carlo, Hausmann, Christiansen, & Randall, 2003; Rodrigues, Ulrich, Mussel, Carlo, & Hewig, 2017). Here, altruism was seen as "voluntary helping motivated primarily by concern for the needs and welfare of another, often induced by sympathy responding and internalized norms/principles consistent with helping others" (Carlo, Randall, & Rothenberg, 2010, p. 273). Following this definition, Rodrigues, Nagowski, Mussel, and Hewig (2018) provided an account further stressing the aspect of "benevolence" in the altruistic act. They propagated

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2020 The Authors. Psychophysiology published by Wiley Periodicals, Inc. on behalf of Society for Psychophysiological Research

the definition of altruistic acts for "an action that is voluntary, intended to benefit another person, driven by this motivation to help the other person to at least 50% (in order to avoid the domination of other motives like public reputation, see e.g., Carlo & Randall, 2002; Rodrigues et al., 2017) and is benevolent, meaning that there is no intention of harming other persons during the process of helping" (Rodrigues, Nagowski, et al., 2018). Following this more narrow definition of altruism, a behavioral phenomenon that is occurring in third-party dictator games, that is, the punishment of proposers making unfair offers (e.g., Brethel-Haurwitz, Stoycos, Cardinale, Huebner, & Marsh, 2016; Henrich et al., 2006), should not be labeled as "altruistic punishment" (Fehr & Gächter, 2002) if the intention to harm the proposer is the driving motive.

Punishment is a costly behavior and has no inherent direct benefits for another person. Based on the aforementioned definition of altruism, this behavior is not altruistic in a more narrow sense. However, it has been argued that altruistic punishment has indirect altruistic side effects by increasing conformity to social norms that increase cooperation (Fehr & Gächter, 2002). But as an inherent benefit for another person is not given and also the benevolence criterion is violated, the narrow definition of altruism given by Rodrigues, Nagowski, et al. (2018) cannot be applied to such a behavior. Giving a more practical example: If one sees someone pushing a victim down on the street, one may have different options to react to it: One could help the victim up (benevolence, direct benefit) or one could beat the aggressor down (no benevolence, no direct benefit, possibly indirect benefit if the aggressor learns not to do this again instead of getting more angry and lets his/her anger go onto less defensive or supported victims). The first action would fulfill the criteria of an altruistic act in a narrow sense as defined by Rodrigues, Nagowski, et al. (2018), but the latter would not. Of course, there might also be a good (or even prosocial) intention behind the punishment that is provided for the aggressor, but as already mentioned by Carlo and Randall (2002) as well as Rodrigues et al. (2017), there are many different motivations and motives to perform prosocial acts and performing a prosocial act or causing a better outcome for society is not a sufficient condition to call the behavior "altruistic."

1.1 | Punishment and compensation in the third-party dictator game

In a third-party dictator game, the participants see two other players: One player (the dictator) divides an amount of money between these two, but the other (the receiver) has no option to react to the division. Then, the participant may act with their own amount of money in order to alter the resulting shares of the two players. Depending on the task, they might have the opportunity to punish the dictator and/or to compensate the receiver. Although punishment behavior has been called "altruistic" by Fehr and Gächter (2002), the question remains, whether this is an appropriate term. In the literature, "altruistic punishment" is still used very often without hesitation, while an altruistic motive or motivation might be in question due to current findings. Recently, Rodrigues, Nagowski, et al. (2018) showed that different personality traits drive punishment and compensation, respectively, if one is able to perform both behavioral options independently in every trial. In this setting, punishment was linked to trait anger, while compensation was linked to trait altruism. This finding shed new light on the link of punishment with anger (e.g., Fehr & Gächter, 2002; Jordan, McAuliffe, & Rand, 2016; Nelissen & Zeelenberg, 2009; Pedersen, Kurzban, & McCullough, 2013; Seip, Van Dijk, & Rotteveel, 2009, 2014). The concept of anger is defined as "the response to interference with our pursuit of a goal we care about. Anger can also be triggered by someone attempting to harm us (physically or psychologically) or someone we care about. In addition to removing the obstacle or stopping the harm, anger often involves the wish to hurt the target" (Ekman & Cordaro, 2011, p. 365). Acting out anger is not a destructive and negative act per se, but can also be used in a constructive manner (Ekman & Cordaro, 2011). Nevertheless, acting out anger is also not to be seen as an altruistic act, as the act of harming someone definitely is not accompanied by benevolence, although a prosocial result might be achieved by enforcing cooperation rules. Following the relation of compensation to trait altruism and the link of trait anger to punishment, we were interested in extending the findings of Rodrigues, Nagowski, et al. (2018) from differences based on personality traits to an active manipulation. Therefore, we chose to use the third-party paradigm in which both punishment and compensation were available in every trial independently of each other and combined it with a state induction of anger, compared to neutral and positive mood induction videos.

1.2 | Movie based induction of states

The induction of states via movie sequences has been one of the most successful induction methods for mood and emotions (Westermann, Spies, Stahl, & Hesse, 1996). Different types of film sequences were used for the induction procedures over the years. One attempt was to use relevant film excerpts and scenes of movies (e.g., Davidson, Ekman, Saron, Senulis, & Friesen, 1990; Schellberg, Besthorn, Klos, & Gasser, 1990). A validated film set for the purpose of inducing emotions and mood including equivalent neutral clips was provided by Hewig et al. (2005) and three of these movies were used to induce the emotions in the present paradigm.

The target emotional state in this paradigm was anger. In order to get control conditions, also neutral and positive mood-inducing video sequences were used. Additionally, we expected a moderating influence of trait anger on the state anger induction. This moderation was expected because of the different personality trait-based "capability" to react to specific state manipulations that are known in many research fields (e.g., trait activation theory: Tett & Burnett, 2003, frontal asymmetry: Coan, Allen, & McKnight, 2006). The perspective of these theories on the trait-state-moderations is a theoretical framework in which situational properties may interact with the personality trait and therefore moderate their validity. They build upon the well-known principle that situations vary in their relevance to any given trait and that trait differences, therefore, will matter to different extents in different situations (e.g., Tett & Guterman, 2000). Combining these trait activation theories with the differential trait-related finding of trait anger leading to more punishment and trait altruism leading to more compensation in a third-party dictator game (see Rodrigues, Nagowski, et al., 2018), we were interested in the state-related activation of the traits and the influence of the states themselves on the third-party dictator game behavioral responses as well as on their influence on the electro-cortical reactions. Hence, we also measured the relevant traits anger and altruism to investigate the trait-state-interactions of anger and altruism on punishment and compensation.

1.3 | Electrophysiological correlates of altruistic behavior

In economic games, psychophysiological markers that were used to link brain activation to behavior are the feedback-related negativity (FRN) and midfrontal theta band activation. These components that have been associated with the evaluation of expectancy (e.g., Hewig et al., 2011; Holroyd & Coles, 2002; Polezzi, Lotto, Daum, Sartori, & Rumiati, 2008) originate in the anterior cingulate cortex (e.g., Debener et al., 2005; Gehring & Willoughby, 2002; Hewig et al., 2007; Miltner, Braun, & Coles, 1997). The component can be measured approximately from 200ms to 400ms after feedback onset (e.g., Holroyd, Pakzad-Vaezi, & Krigolson, 2008) and has a negative deflection, leading to a more negative amplitude if the outcome is worse than expected (Holroyd & Coles, 2002). More recent research, however, came to a different perspective concerning this event-related component and introduced the term reward positivity (Rew-P; Baker & Holroyd, 2011) instead of FRN. Hajcak, Moser, Holroyd, and Simons (2006) found this Rew-P to be linked to positive feedback and repressed by negative feedback (Proudfit, 2015). Holroyd et al. (2008) as well as Baker and Holroyd (2011) could additionally show that the reactivity to reward is driving differences in the component complex formally named FRN (see also Hewig et al., 2007). Having identified the dependence of this electrophysiological component on reward, yet the importance of the negative feedback for negative amplitudes of the component was still found (e.g., Hajcak et al., 2006; Hewig et al., 2007). However, the proposed mechanism leading to this negative deflection was changed from the negative violation of expectancy (Holroyd & Coles, 2002) to a global outcome evaluation process (Kujawa, Smith, Luhmann, & Hajcak, 2013), leading to a higher FRN if reward and therefore the Rew-P is absent.

In economic games, the FRN was not only used for behavioral outcome evaluation, but also for the evaluation of offers that are given. Boksem and De Cremer (2010) showed that the context of the perception of fairness was associated with the magnitude of the FRN, with higher FRNs being recorded to unfair offers if fairness was important to the participant. Similar findings arose for an active manipulation of the perceived fairness of agents with previous social exclusion games (Qu, Wang, & Huang, 2013), leading to enhanced FRN for unfair players. This enhanced FRN for unfair offers was also found by Mothes, Enge, and Strobel (2016) as well as Sun, Tan, Cheng, Chen, and Qu (2015) in the third-party dictator game. These findings further illustrate the modulation of the FRN by the perception of fairness, although the participants are not directly affected by the offer as in an ultimatum game. Boksem and De Cremer (2010) also argued that in bargaining situations the FRN might not be driven primarily by the outcome of the participants, but by the violation of the fairness norm. This claim was bolstered by the findings of Mothes et al. (2016) as well as from Sun et al. (2015), as their participants' monetary outcome was not linked to the offers in their paradigm, yet they showed higher FRNs to unfair offers. This stresses the evaluative nature of the FRN component, as the FRN can be seen independently of an evaluation of direct personal outcome. But the FRN could also be used as a predictor for behavior. Cohen and Ranganath (2007) showed that it is possible to predict subsequent binary decisions of playing participants in a strategic game on basis of the FRN or theta activation. In more complex paradigms based on free choice however, the FRN may have had some limits due to being derived from trial means. In order to also use the electrophysiological correlates as a predictor in the dictator game for subsequent behavioral decisions, midfrontal theta activation has been used to parameterize the FRN reaction on basis of single trial (Cavanagh, Cohen, & Allen, 2009; Cavanagh, Zambrano-Vazquez, & Allen, 2012; Cohen, Elger, & Ranganath, 2007; Rodrigues, Ulrich, & Hewig, 2015). This frequency approach can be used as a reliable measure in single-trial contexts and is, therefore, an appropriate index for free choice paradigms to predict behavioral responses. Accordingly, for more complex decisions in the dictator game context, Rodrigues et al. (2015) found that midfrontal theta predicted fair offers in the dictator game in participants with high trait altruism.

As we were using a third-party dictator game in this study with complex non-binary behavioral options, we had the interesting possibility to disentangle the perception of an offer as unfair from the actual behavioral response of the participants, as they are not personally affected by the fairness of the offer. Thus, a difference in the FRN compared to midfrontal theta band activation could be possible in the third-party dictator game: The FRN might be influenced predominantly by the fairness of the offer, while midfrontal theta is more strongly linked to the resulting behavior and may also reflect the necessary cognitive control to execute this behavior (e.g., Cavanagh et al., 2009; Cavanagh & Frank, 2014; Cohen, 2011; Phillips, Vinck, Everling, & Womelsdorf, 2014; Polanía, Nitsche, Korman, Batsikadze, & Paulus, 2012). As explained above, we would expect the FRN to be an evaluative signal concerning the expectancy and perception of fairness (compare Holroyd & Coles, 2002), while the midfrontal theta band activation may be indicative of altruistic behavioral responses (e.g., Rodrigues et al., 2015 for dictator game offers) or the cognitive control (e.g., Cavanagh et al., 2009; Cavanagh & Frank, 2014; Cohen, 2011; Phillips et al., 2014; Polanía et al., 2012) that is necessary to execute this behavior (e.g., Mussel, Göritz, & Hewig, 2013).

But behavioral decisions in economic games are not only dependent on the perception of offer and situation, but also on relevant personality traits (e.g., Rodrigues et al., 2015). Hence, it is also interesting to investigate the relation of both midfrontal theta and the FRN with personality traits, especially in the third-party dictator game. As mentioned above, Rodrigues et al. (2015) found higher midfrontal theta for high trait altruism while making fair offers in the dictator game, indicating a higher midfrontal theta band activation if an altruistic act is performed. They argue that this midfrontal theta band activation is related to midcingulate cortex and posterior cingulate cortex activation (see Cohen, 2011; Weiland, Hewig, Hecht, Mussel, & Miltner, 2012), that indicates a moral sensitivity or justice sensitivity. As the midcingulate cortex, as well as the insula, was generally found to be involved in altruistic behavior and empathy (Mathur, Harada, Lipke, & Chiao, 2010), the neuronal basis of altruistic acts seen in the dictator game are moderated by trait altruism, enhancing the midcingulate activation which can be measured via theta band in EEG.

Concerning the outcome evaluation, the FRN reaction that was shown to be linked to unfair offers (e.g., in the ultimatum game: Boksem & De Cremer, 2010; Hewig et al., 2011; Polezzi, Daum, et al., 2008) has also been shown to be moderated by relevant traits. The underlying neural structures of the FRN reaction are considered to be the anterior cingulate cortex (Holroyd & Coles, 2002) and are also found to be involved in the neural reaction to unfair offers in the ultimatum game (e.g., Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003), Moreover, Wu, Leliveld, and Zhou (2011) could show that if no reaction is involved, also the FRN reaction is present to unfair offers and this reaction was also moderated, in this case by the closeness of the proposing dictator to the receiver. Boksem and De Cremer (2010) additionally showed the moderation of the FRN by the concern for fairness. They argue that the importance of social or moral norms relate to the magnitude of the FRN, because the subjective importance of following or violating these norms of society cause different magnitudes of anterior cingulate cortex reaction. Concerning the task that we are about to use, Mothes et al. (2016), as well as Sun et al. (2015), found that the FRN for unfair offers in a third-party dictator game is modulated by trait altruism or trait empathy, which is closely related to altruism (Leliveld, Dijk, & Beest, 2012). Sun et al. (2015) found higher FRN for high altruism while Mothes et al. (2016) found lower FRN for high empathy. Mothes et al. (2016) argue that the anterior cingulate cortex activation is moderated by the sensitivity to fairness norm violations as Boksem and De Cremer (2010) could show before. They explain their results by proposing that unfair offers are much more unexpected if one is not able to be empathic, while Sun et al. (2015) argue that the higher FRN is linked to a lower expectancy of unfair offers in persons with high altruism. But putting together both studies, they both stress that the expectancy and importance of fairness is a key element of the resulting FRN and its modulation by traits.

Independent of this fairness expectancy driven FRN modulation, Angus, Kemkes, Schutter, and Harmon-Jones (2015) showed that anger leads to higher FRN in a gambling task. They explain this finding with the motivational properties of anger, being an approach-related emotional state. As participants that value the reward of the gambling task are deprived of their goal that they wanted to achieve, they show an enhanced FRN. Hence, anger may lead to an amplification of the FRN, if an inherent goal (or in our case: a norm, reflected by the fairness of an offer) is not reinforced but non-rewarded or the respective norm is violated.

Summing up the findings about the electrophysiological correlates of unfair offers and related behavioral responses, the FRN may be an indicator of a personality-related evaluation of a situation or offer (e.g., Mussel, Reiter, Osinsky, & Hewig, 2015), while midfrontal theta band activation may reflect a personality-related amount of cognitive control (e.g., Mussel et al., 2013), resource investment (Mussel, Ulrich, Allen, Osinsky, & Hewig, 2016) or behavioral decision process (Cohen & Ranganath, 2007; Rodrigues et al., 2015). Importantly, Mussel et al. (2015, 2016) implemented a state-trait interaction perspective on personality using a trait-inducing situation and showed that this did indeed strongly influence the behavioral and electrophysiological results accordingly.

1.4 | Hypotheses

Integrating the recent findings about the relation of trait anger with punishment and trait altruism with compensation in the third-party dictator game as well as the findings of the evaluative character of the FRN and the predictive value of midfrontal theta activation in EEG with the state induction, we had following hypotheses:

On the behavioral level, we expected an influence of state induction on punishment behavior, with higher levels of punishment in the anger condition compared to any other condition. Furthermore, we expected that this effect is moderated by trait anger, with even higher punishment in the anger condition compared to any other condition for high levels of trait anger compared to low levels of trait anger.

On the electrophysiological level, we expected a higher FRN for unfair offers compared to fair offers of the dictator. This effect should be more pronounced during the anger state induction compared to any other condition. Additionally, this effect should be moderated by trait anger, leading to a stronger effect for individuals with high compared to low levels of trait anger. For trait altruism, we expected a FRN modulation.

Also, we hypothesized that midfrontal theta band activation measured at the presentation of a dictator offer is positively related to compensation. For punishment, an inverse link was predicted. Also, we expected the theta band activation to be moderated by trait altruism and trait anger, with higher midfrontal theta in high compared to low altruistic participants and lower midfrontal theta band activation in participants with high trait anger compared to low trait anger.

2 | MATERIALS AND METHOD

2.1 | Ethical statement

The study was carried out in accordance with the recommendations of "Ethical guidelines, The Association of German Professional Psychologists" ("Berufsethische Richtlinien, Berufsverband Deutscher Psychologinnen und Psychologen") with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki before they participated in the experiment. The protocol was approved by the local ethics committee of the department of psychology of the Julius-Maximilians-University of Würzburg (Ethikkommission des Institutes für Psychologie der Humanwissenschaftlichen Fakultät der Julius-Maximilians-Universität Würzburg).

2.2 | Participants

We a priori estimated the required sample size with G-power software (Faul, Erdfelder, Lang, & Buchner, 2007). Assuming an average effect of $r = .36$ of anger on altruistic punishment (e.g., Lotz Baumert, Schlösser, Gresser, & Fechtenhauer, 2011), $\alpha = .05$ and power $(1 - \beta) = .8$ yielded a required sample size of $N = 55$. Sixty-four persons participated in this study. After raw data inspection, 7 datasets had to be rejected because of poor data quality. Additionally, one participant had to be excluded because of the lack of trait recordings. Hence, 56 participants were included in the analysis (14 male, mean age = 26.16, $SD = 8.34$, range = 19–60). All participants were right-handed, had normal or corrected to normal vision and most of them were students. The participants were paid 15€ or they received course credits for their participation.

2.3 | Paradigm

The employed paradigm was a third-party economic game, in which participants were observing a dictator game (e.g., Güth, 1995). Dictators divided an amount of money between themselves and receivers. The receivers had no option to react to the offer. Participants were able to perform compensation to receivers and punishment to dictators in every trial independently. They saw different offers from different dictators to different receivers in every trial of every block. Unbeknownst to the participant, the dictator games were simulated without real participants. They were debriefed about this deception after the experiment. All trials started with the alleged offer of a fictive dictator shown for 1.5 s illustrated as picture of an offer with either 8:0, 6:2 or 4:4 credits (1 credit = 1 Cent) and therefore always leaving at least one half of the money for the fictive dictators. Then participants had the opportunity to spend their money on compensating the receiver and punishing the dictator (5 s for each decision). This time constraint was imposed because of the cover story and to keep the experiment time under control, for the free choice time could lead to very long trials. The amount of money participants were able to spend on each action was identical to the money kept by the dictator. It was given to them in each trial anew. This amount of money was indicated to the participant via the maximum in the decision displays (see also example in Figure 1). For example, if dictators kept 8 credits for themselves a maximum of 8 credits could be spent on compensation, but also 8 credits could be used on punishment. Hence as the money to spend on each behavior was 8 credits in this example the participants got a total amount of 16 credits to spend. If the dictator kept only 6 credits in a trial (see example Figure 1), then the decision displays maximum was not 8 credits per decision (leading to 16 credits in total for the trial) but only 6 credits per decision (leading to 12 credits in total for the trial). All money that was not spent by the participants during a trial was kept by them. Therefore, a participant could for instance use the maximum amount for one behavior (half of the total amount of the trial) for punishment and the dictator would get 0 credits in any trial. Concerning the compensation, if a participant used the maximum amount for one behavior, the receiver would get 8

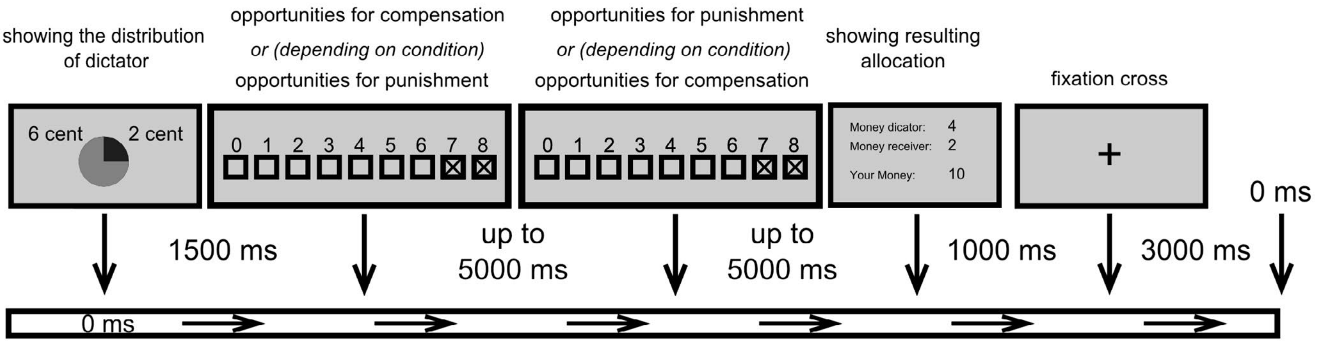

FIGURE 1 Schematic display of a trial. Note that the participants can only spend as much money as the dictator offers on each behavioral option (compensation and punishment). Hence, the participants get to spent 2 × 6 credits = 12 credits in this example in total (1 credit = 1 Cent). In this example, 2 credits were spent on punishment and 0 credits on compensation. The participants experienced 72 Trial per block. There were three blocks with an emotion induction video sequence before every block (see Figure 2)

credits in any trial. Thus, the resulting amount of money for dictator and receiver were kept between the same boundaries for every trial, being 0 credits for the dictator and 8 credits for the receiver if all available money was spent by the participant for the respective behaviors. We only analyzed the relative amount of money used, meaning the absolute amount of money that was spent by the participant, divided by the amount of money that was available in the respective trial, in order to correct for the different reference frames. For example, 8 credits may be less worth if one has 16 credits to spend in total during the trial versus. 8 credits to spend in total. The order of the decisions was counterbalanced between participants, with compensation being first for half of the participants and punishment being first for the other half (odd/even split). After having made both decisions or after the time had passed (5 s for each action) the trial continued with a display of the resulting allocation for the three parties for one second. Thereafter, a fixation cross was shown for 3 s (Figure 1). Each block of the paradigm consisted of 72 trials (24 trials of each offer by the dictator). Before every block, a state inducing video was shown (cp. Hewig et al., 2005). The videos were either neutral (*Crimes and Misdemeanors*), anger inducing (*My Bodyguard*) or positive (*An Officer and Gentleman*). The order of the videos was counterbalanced between participants (see Figure 2).

It was stated clearly to the participants that the offers displayed to them were made by two other players that had played the dictator game in a previous session. It was also stated, that these two players would be paid after the experiment together with the participant, depending on their past decisions and those of the participant. This cover story was reinforced by collecting the bank account number of every participant via written consent at the beginning of the experiment on a separate piece of paper. During the experiment, the bank account data was still lying on the participants' table. After the experiment, the participants were debriefed about the deception, compensated for their participation and then reminded to take their bank account information with them.

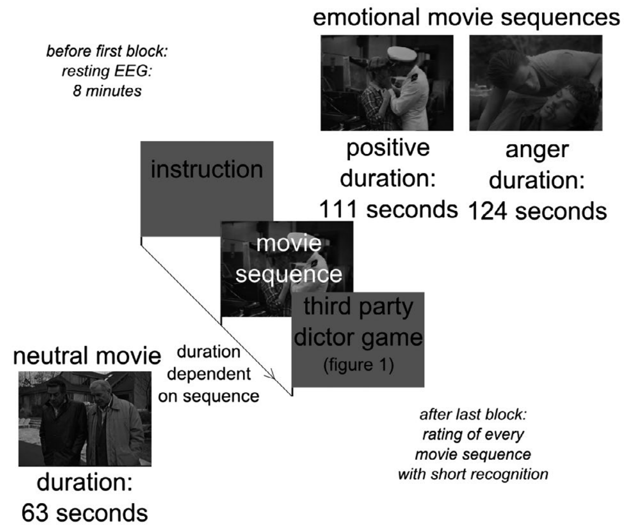

FIGURE 2 Procedure of the experiment. Three different video sequences were used for emotion induction. The order of the video sequences was counterbalanced. The movie sequences were taken from Hewig et al. (2005). Note that the resting EEG is not included in the scope of the manuscript and only mentioned in order to give a complete overview of the procedure

2.4 | Procedure

Before coming to the laboratory, participants filled in a web-based questionnaire to assess the relevant traits. Also, demographical data were collected (gender, age and handedness). The online questionnaire was presented with SoSci Survey (Leiner, 2016), an online questionnaire platform.

At the beginning of the laboratory session, information material about the session and the informed consent form was given to the participants. Then, participants were seated in front of a 61 cm (24″) widescreen monitor in 50–60 cm distance and an EEG cap were placed on their head, as well as electrodes for skin conductance on their left hand and electrodes for the heart rate on their collarbones and the left costal arch. Additionally, headphones were placed on the head of the participants in order to provide tones during the paradigms and instructions for an 8 min resting EEG period that is not included in the scope of this manuscript.

After the resting EEG period, participants played the third-party economic game, which consisted of three blocks with 72 trials each. Before every block, a neutral, anger inducing, or positive video was shown. The three different stimuli were selected because of the target emotion (anger): As anger is an emotion that can be characterized by having an approach motivation but not a positive valence (e.g., Harmon-Jones & Allen, 1998; Harmon-Jones, Sigelman, Bohlig, & Harmon-Jones, 2003; Hewig, Hagemann, Seifert, Naumann, & Bartussek, 2004; Rodrigues, Müller, Mühlberger, & Hewig, 2018), we wanted to incorporate both a neutral comparison (valence control) and one that incorporates approach motivation as can be found in positive emotional states (motivation control). The order of the videos was counterbalanced. At the end of the session, participants were asked about the movies concerning their impressions and feelings during the presentation of the film (see Ratings section below). Then they were debriefed about the deception of seeing offers of other players in the dictator game, given back their bank account information as stated in 2.3 and then they were freed of the apparatus and given the compensation for the session.

2.5 | Apparatus

2.5.1 | EEG recording

The EEG was measured by Ag/AgCl-electrodes located in an electrode cap in the following 32 positions: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Pz, FC1, FC2, CP1, CP2, FC5, FC6, F9, F10, TP9, TP10, PO9, PO10, FCz, and Cz (according to the international 10–10 system). Ground electrode was located on AFz position, the reference electrode was Cz. An additional electrode to register eye movements and blinks was put below the left eye.

Electrode impedances were kept below 5 kOhm for the EEG. Data were recorded with a sampling rate of 250 Hz and a bandpass filter from form 0.3–70 Hz with BrainVision BrainAmp Standard and Brain Vision Recorder software. For further computation, MATLAB and EEGLAB toolbox (Delorme & Makeig, 2004) were used. Raw data were inspected automatically using the z-value-based channel rejection for probability, kurtosis and frequency range from 1–125 Hz, using the outlier criterion of $z = 3.29$ (see Tabachnick & Fidell, 2014) and bad channels were spherical interpolated automatically. After the raw data segmentation of the events, ranging from −1.5 to 2.5 s around every dictator offer, the data was filtered with a 1–40 Hz Butterworth bandpass filter. A first independent component analysis (ICA; Makeig, Debner, Onton, & Delorme, 2004) was computed and bad segments were rejected using IC-based z-value artifact detection with the criterion of $z = 3.29$ (see Tabachnick & Fidell, 2014) for probability and kurtosis. After this rejection of segments, a second ICA was performed and the addons ADJUST (Mognon, Jovicich, Bruzzone, & Buiatti, 2011) and MARA (Winkler, Haufe, & Tangermann, 2011) were used to automatically detect and reject artifact components. Then, CSD transformation was performed using the CSDToolbox (Kayser & Tenke, 2006) and the data was segmented from −1 to 2 s around every offer of the dictator with a baseline from −0.5 to 0 s. For theta band activation, wavelets extraction from 3.5 to 8.5 Hz was used with log spacing and fixed cycles (4.5). The wavelet extraction script was based upon code provided by Cohen (2014).

FRN and theta quantification

For the quantification of the FRN and theta band activation, the electrode position FCz was used after visual inspection of the topographical trial time course. For peak-analysis window selection, automatic peak detection was performed on the grand means during the time window of 200–300 ms (Yeung & Sanfey, 2004), leading to a peak at 280 ms for the FRN and a peak at 244 ms for the theta band activation. An analysis window of 40 ms was used around the peaks from 260 ms to 300 ms for FRN and from 224 ms to 264 ms for midfrontal theta band activation. Cronbach's α of the FRN and midfrontal theta band activation can be seen in Table 1.

2.6 | Ratings

The participants were provided with questions about the conditions and paradigms, assessing the concepts of negative emotions (excluding questions about anger), positive emotions and anger during the different movie conditions with 21 items on a scale from 0–9 (0 indicated *not at all* and 9 indicated *very strong*; see Hewig et al., 2005). These questions were asked in order to check the manipulation via video sequences and the scales were condensed into scales of negative emotions, positive emotions, and anger in order to differentiate between negative emotions per se and anger, as well as positive emotions that may specifically be targeted by the positive emotion induction. The results of the manipulation check can be seen in Supporting Information.

2.7 | Trait measurement

The traits were assessed via SoSci Survey online-questionnaire-portal (Leiner, 2016). For prosocial tendencies, the German version of the revised version of the prosocial tendency measure (PTM-R, Carlo & Randall, 2002, Rodrigues

TABLE 1 Cronbach's alpha for the electrophysiological signals in the different conditions

Electrophysiological signal	Movie	Dictator keeping: 4 credits	Dictator keeping: 6 credits	Dictator keeping: 8 credits
FRN	Neutral movie	.864	.787	.895
FRN	Happy movie	.791	.754	.541
FRN	Anger movie	.814	.782	.812
ln midfrontal theta band activation	Neutral movie	.943	.964	.966
ln midfrontal theta band activation	Happy movie	.973	.975	.984
ln midfrontal theta band activation	Anger movie	.967	.952	.963

et al., 2017, Cronbach's $\alpha_{altruism}$ = .621) was used. For measuring trait anger the trait items of the German version of the State- trait- anger-expression-inventory (STAXI; Schwenkmezger & Hodapp, 1991; Spielberger, 1988, Cronbach's α_{anger} = .783) was used.

2.8 | Statistics

2.8.1 | Behavior & EEG

The behavioral reactions, as well as the FRN and midfrontal theta band activation on electrode position FCz, were analyzed using 3 × 3 × 2 × 3 general linear models with the within factors movie (anger/happy/neutral) and money kept by dictator (8/6/4), the between factors order of behavior (punishment first/ compensation first) and position of the anger movie (first/second/third) with the z-standardized continuous predictors trait altruism and trait anger. The traits were included because of the finding of Rodrigues, Nagowski, et al. (2018), where punishment and compensation were predicted by these variables. Two separate general linear models were computed for the dependent variables compensation and punishment, respectively, as well as one for the FRN and one for the midfrontal theta band activation. *P*-values of subsequent tests to disentangle main effects and interactions were Bonferroni-Holm corrected. Greenhouse-Geisser correction factors can be seen in Table S1 in supporting Information.

Midfrontal theta and the FRN were used as a predictor for subsequent behavior. Hence, models for the criterion relative punishment (credits used for punishment/ credits available for punishment) and relative compensation (credits used for compensation/ credits available for compensation) were computed including either midfrontal theta activation or the FRN as predictor and therefore leading to four different resulting models. The proposed multilevel models had the level 1 predictors Ln(theta) or FRN, movie (anger/happy/neutral), order of behavior (punishment first/compensation first), position of anger movie and money kept by dictator (8/6/4). The electrophysiological signals were centered for every participant on the trial level (level 1) and the grand mean centered mean participant electrophysiological signal was added on level 2 together with the relevant traits altruism and anger. The inclusion of the electrophysiological signal on level 2 and on level 1 was done in order to address the between effects as well as the within participant variations (West, Ryu, Kwok, & Cham, 2011).

All statistical analyses were carried out using R software (R Core Team, 2019) with the packages "afex" (Singmann, Bolker, Westfall, & Aust, 2019) with "lsmeans" (Lenth, 2016) and "nlme" (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2018) and SPSS (Version 23). For graphical illustration, the package "ggplot2" (Wickham, 2016) with "ggstance" (Henry, Wickham, & Chang, 2018) and "ggsignif" (Ahlmann-Eltze, 2019) was used. For reading in data in R, "readxl" (Wickham & Bryan, 2019) was used.

3 | RESULTS

3.1 | Behavior

For punishment, a main effect of money kept by the dictator was significant $F(2,96) = 30.33$, $p < .001$, $\eta^2_p = .39$ leading to more punishment the more money was kept by the dictator (*t*s > 3.26, *p*s < .01, *d*s > .331, see Figure 4a). Also, the interaction of the position of the anger movie with the type of movie was significant $F(4,96) = 5.81$, $p < .01$, $\eta^2_p = .19$, with higher punishment after the anger movie if it was shown first to the participants (*t*s > 3.58, *p*s < .01 *d*s > .689, see Figure 3b). In all other conditions, this difference was not significant (*t*s < 1.25, *p*s > .67). Additionally, higher trait anger led to higher punishment $F(1,48) = 4.29$, $p < .05$, $\eta^2_p = .08$, $r = .11$. Additional marginal effects can be seen in Table S2 in the Supporting Information.

For compensation, a main effect of the money kept by the dictator was also significant $F(2,96) = 56.29$, $p < .001$, $\eta^2_p = .54$ leading to more compensation the more money was kept by the dictator (*t*s > 3.30, *p*s < .01, *d*s > .335, see Figure 4a). This effect was modulated by trait anger $F(2,96) = 5.38$, $p < .05$, $\eta^2_p = .10$. Subsequent tests of the

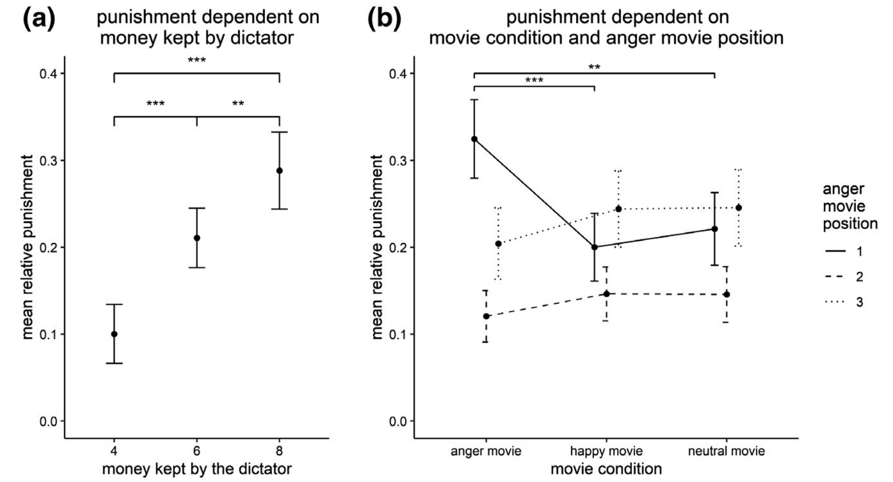

FIGURE 3 Punishment dependent on money kept by the dictator and the interaction of the movie sequence and the anger movie sequence position. Error-bars represent mean *SE*

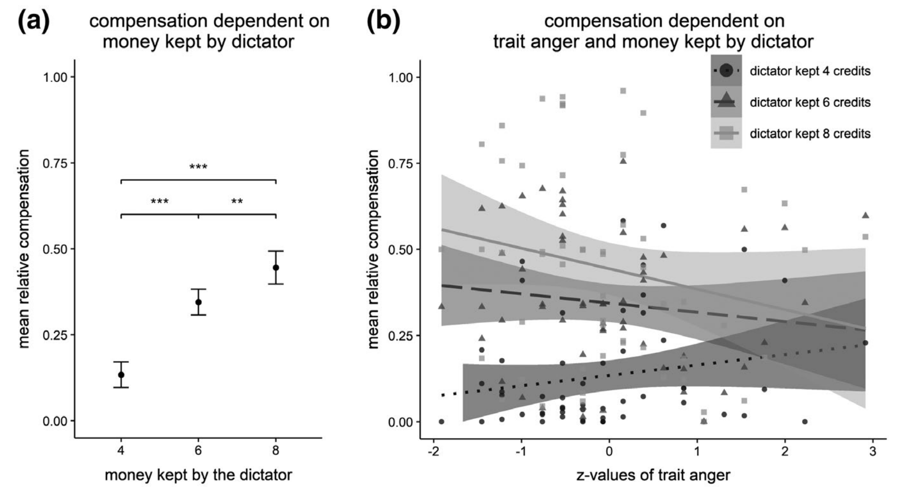

FIGURE 4 Compensation dependent on the money kept by the dictator and the trait anger of the participant. Error-bars represent mean *SE*. Shaded areas represent the 95% confidence-interval

correlation of trait anger and the compensation during the paradigms revealed a significant difference for the dictator keeping only 4 credits ($r_{dictator\ kept\ 4}$ = .186) to the dictator keeping more money (*z*s > 2.21, *p*s < .05, $r_{dictator\ kept\ 8}$ = −.211, $r_{dictator\ kept\ 6}$ = −.132, $z_{dictator\ kept\ 6\ vs\ 8}$ = −1.40, *p* = .08, with the relation being more negatively if the dictator kept more money (Figure 4b).

Additional marginal effects can be seen in Table S2 in the Supporting Information.

Summing up the behavioral effects, both compensation and punishment were higher, if more credits were kept by the dictator. Punishment was higher if the anger movie was presented as the first movie. Also, compensation was lower if trait anger was high and the dictator kept many credits.

3.2 | FRN on FCz

For the FRN on electrode position FCz, no significant main effect could be found (*F*s < 1.68, *p*s > .20) and only a marginal effect could be identified for the money kept by the dictator $F(2,96) = 2.52$, $p = .09$, $\eta_p^2 = .05$ (see Figure 5). However, a significant interaction of the money kept by the dictator and trait anger could be identified $F(2,96) = 3.47$, $p < .05$, $\eta_p^2 = .07$. Subsequent correlation analyses revealed a more negative relation of the FRN and trait anger (meaning more pronounced FRN amplitudes for higher trait anger) if the dictator kept 8 credits ($r_{dictator\ kept\ 8}$ = −.213) compared to the conditions if the dictator kept less credits (*z*s = −1.77, *p* < .05, $r_{dictator\ kept\ 4}$ = −.064, $r_{dictator\ kept\ 6}$ = .016, $z_{dictator\ kept\ 6\ vs\ 8}$ = 0.84, *p* = .201, see Figure 6).

Also the interaction of the position of the anger movie and the money kept by the dictator was significant $F(4,96) = 2.83$, $p < .05$, $\eta_p^2 = .11$). Subsequent Bonferroni-holm adjusted analyses revealed that the FRN was more negative if the dictator was keeping 6 credits than for keeping 4 or 8 credits if the anger movie was on second position (*t*s > 2.74, *p*s < .05, *d*s > .45).

Other marginal effects can be seen in Supporting Information in Table S2.

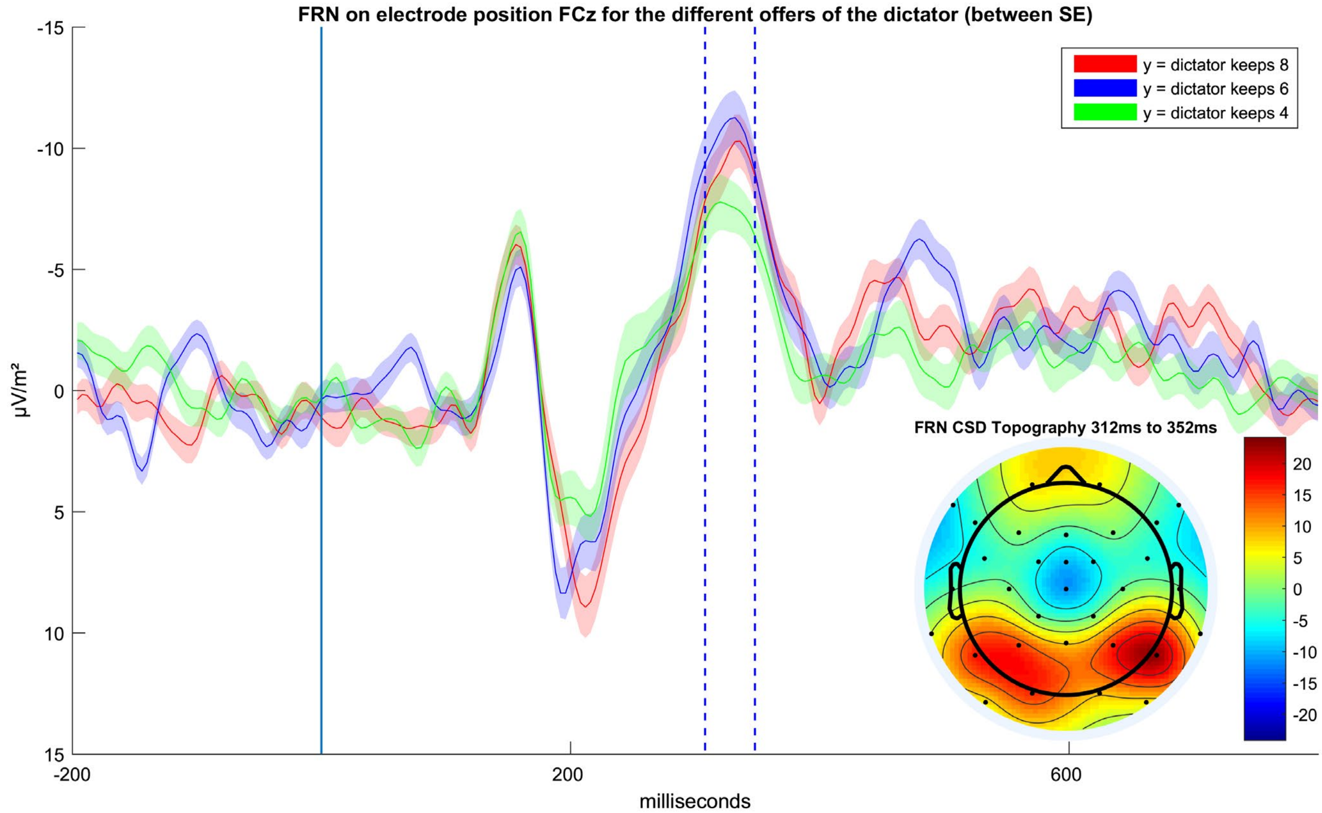

FIGURE 5 Current source density signal on electrode position FCz around the dictator offer and topography of the EEG-signal during automatically detected time window. Note that the dashed lines mark the automatically detected FRN time window

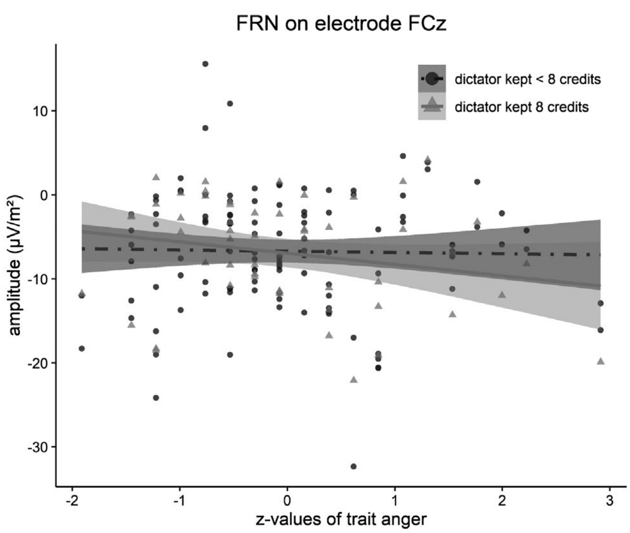

FIGURE 6 Feedback related negativity on electrode position FCz dependent on the money kept by the dictator and the trait anger of the participant. Shaded areas represent the 95% confidence-interval

Summing up the effects of the FRN analyses, higher trait anger led to more negative FRN values if the dictator kept 8 credits. Also, if the anger movie was on second position and the dictator kept 6 credits, the FRN was more negative.

3.3 | Single-trial FRN on FCz as a predictor of behavior

The single-trial FRN was used as a predictor of behavior in a multilevel model predicting the behavior on a single trial basis. The model fit determined with corrected Akaike criterion (AICC) was best for the alternative model 1 for relative compensation as well as for relative punishment (see Table S3 in Supporting Information). Due to the model fit criteria and the simpler fixed effect structure, we selected the alternative model 1 as the fitting results.

The model for punishment behavior revealed a significant fixed effect for trait anger, leading to more punishment if the trait anger was higher (Figure 7, Table 2). Integrating the main effects of the money kept by the dictator, the punishment was higher if 6 or 8 credits were offered (Figure 3, Table 2). Integrating the main effects and interactions of the type of movie and the anger movie position, the earlier the anger movie was presented, the higher the punishment, while this effect was dampened for other movie types (Figure 3, Table 2).

Additional marginal effects can be seen in Table S5 in the Supporting Information.

For compensation, the model revealed only fixed effects for the money kept by the dictator, with higher compensation if the dictator kept more money (Figure 4a, Table 3).

Summing up the effects of the single trial FRN analyses, higher trait anger led to more punishment behavior. Compensation and punishment were higher, if more credits were kept by the dictator and punishment was higher if the anger movie was presented as the first movie.

3.4 | Theta band activation on FCz

For the theta band activation on electrode position FCz (see Figure 8), a significant main effect for the money kept by the dictator could be found $F(2,98) = 23.91$, $p < .001$, $\eta_p^2 = .33$, leading to more midfrontal theta activation if the dictator kept only 4 credits (ts > 5.42, ps < .001, ds > .545). Also, a marginal effect could be identified for the interaction of money kept by the dictator and the movie $F(4,196) = 2.16$, $p = .08$, $\eta_p^2 = .04$, with the previous effect of higher theta activation for 4 credits being only true for anger (ts > 4.94, ps < .001, ds > .289) and partly for happy ($t_{\text{dictator kept 4 vs. 8}}(290) = 4.49$, ps < .001, $d = .263$).

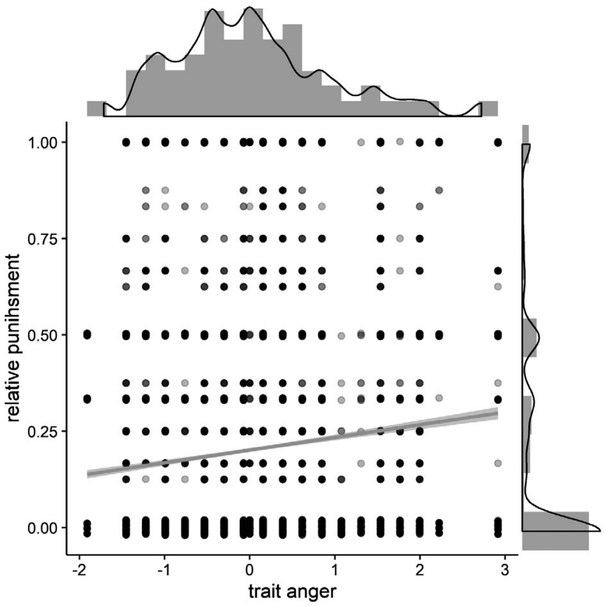

FIGURE 7 Relative punishment depending on trait anger

3.5 | Single trial theta band activation on FCz as a predictor of behavior

The theta band activation [3.5–8.5 Hz] was used as a predictor of behavior in a multilevel model predicting the behavior on a single trial basis. The model fit determined with the corrected Akaike criterion (AICC) was best for the alternative model 1 for relative compensation as well as for relative punishment (see Table S3 in Supporting Information). Due to the model fit criteria and the simpler fixed effect structure, we selected the alternative model 1 as the fitting results.

The model for the punishment behavior revealed a significant fixed effect for mid-frontal theta band activation during the trials with more mid-frontal theta activation leading to less punishment behavior (see Figure 9, Table 4). Also participants with more mean midfrontal theta compared to other participants punished less (Table 4). Integrating the main effects of the money kept by the dictator, the punishment was higher if 6 or 8 credits were offered (Figure 3). Integrating the main effects and interactions of the type of movie and the anger movie position, the earlier the anger movie was

TABLE 3 Intercept and significant fixed effects for relative compensation single trial model with FRN as predictor for behavior

Fixed effects	Value	SE	*df*	*t*-value	*p*-value
dictator kept 6 credits	0.249	0.084	12,186	2.973	.003
dictator kept 8 credits	0.366	0.113	12,186	3.240	.001

Note: Baseline category: dictator kept 4 credits and the anger movie was watched at the beginning of the block.

TABLE 2 Intercept and significant fixed effects for relative punishment single-trial model with FRN as a predictor for behavior

Fixed effects	Value	*SE*	*df*	*t*-value	*p*-value
(Intercept)	0.200	0.054	12,186	3.678	.000
Dictator kept 6 credits	0.159	0.060	12,186	2.640	.008
Dictator kept 8 credits	0.233	0.097	12,186	2.407	.016
Trait anger	0.034	0.014	52	2.492	.016
Happy movie	−0.167	0.049	12,186	−3.408	.001
Neutral movie	−0.132	0.060	12,186	−2.195	.028
Position of the anger movie × happy movie	0.080	0.023	12,186	3.513	.000
Position of the anger movie × neutral movie	0.067	0.028	12,186	2.414	.016

Note: Baseline category: dictator kept 4 credits and the anger movie was watched at the beginning of the block.

FIGURE 8 Frequency spectrum on electrode position FCz around the dictator offer and topography of the theta band activation during automatically detected time window

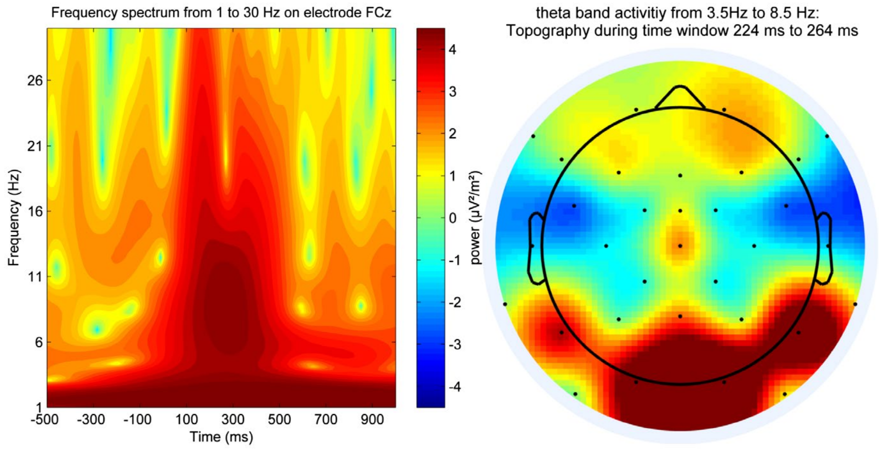

FIGURE 9 Relative punishment dependent on the participant-centered theta band activation and the grand mean centered mean theta band activation of the participant on electrode position FCz. Shaded areas represent the 95% confidence-interval

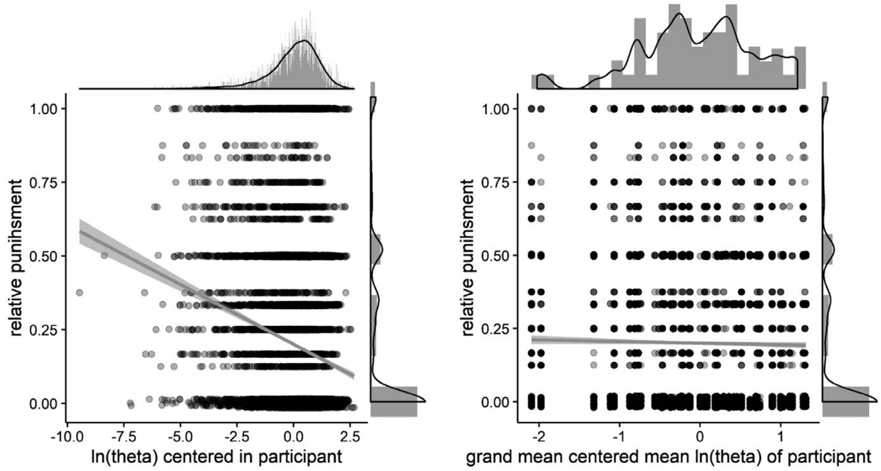

TABLE 4 Intercept and significant fixed effects for relative punishment single-trial model with theta band activation as predictor for behavior

Fixed effects	Value	*SE*	*df*	*t*-value	*p*-value
(Intercept)	0.240	0.057	12,186	4.200	.000
Dictator kept 6 credits	0.156	0.059	12,186	2.616	.009
Dictator kept 8 credits	0.240	0.094	12,186	2.548	.011
Ln(theta) on FCz (centered in participant)	−0.030	0.005	12,186	−6.542	.000
Mean Ln(theta) of participant	−0.056	0.017	52	−3.209	.002
Happy movie	−0.201	0.047	12,186	−4.242	.000
Neutral movie	−0.128	0.061	12,186	−2.118	.034
Position of the anger movie	−0.063	0.026	52	−2.378	.021
Position of the anger movie × happy movie	0.096	0.022	12,186	4.390	.000
Position of the anger movie × neutral movie	0.065	0.028	12,186	2.336	.020

Note: Baseline category: dictator kept 4 credits and the anger movie was watched at the beginning of the block.

presented, the higher the punishment, while this effect was dampened for other movie types (Figure 3). Accordingly, the latter two effects were not influenced by theta activation.

For compensation, the model revealed only fixed effects for the money kept by the dictator, with higher compensation if the dictator kept more money (Figure 4a, Table 5).

TABLE 5 Significant fixed effects for relative compensation single-trial model of theta band activation

Fixed effects	Value	*SE*	*df*	*t*-value	*p*-value
Dictator kept 6 credits	0.326	0.070	12,186	4.661	.000
Dictator kept 8 credits	0.464	0.088	12,186	5.283	.000

Note: Baseline category: dictator kept 4 credits and the anger movie was watched at the beginning of the block.

In summary, the single-trial midfrontal theta analysis additionally revealed that more theta activation was related to less punishment behavior both intraindividually and interindividually.

4 | DISCUSSION

In this study, we found that punishment behavior was influenced by the money kept by the dictator and the anger induction as hypothesized. Contrary to our hypothesis, we did not find a moderation of this effect by trait anger, but a direct influence of trait anger on punishment. Also, we found a main effect of money kept by dictator and its interaction with trait anger on compensation. On the electrophysiological level, the hypothesized higher FRN effect for unfair offers was only marginally significant, and the effect was mainly driven by the fair offers having a substantially lower FRN. This is in line with the finding that the FRN or Rew-P is dependent on the best outcome of the trials and not the best outcome that is available at the moment (Kujawa et al., 2013). Also, a significant interaction with the money kept by the dictator and trait anger was found as presumed, although the postulated modulation for trait altruism was not detected. Higher midfrontal theta band activation was found to be related to less punishment as hypothesized, although we did not find more compensation or any trait moderation of this effect that has been suggested.

The anger induction we tried to achieve via the movie sequences was successful, as one may see from the ratings concerning the movie sequences. However, we only found the predicted effect of higher punishment behavior if the anger movie was presented as the first movie. This unexpected result may be due to the time course of the experiment and the ongoing repetition of the decisions that tend to lead to a decision strategy if repeated long enough (e.g., Andreoni, 1988). As can be seen in Figure 3, punishment after the anger movie is exceptionally high if the anger movie is the first movie sequence, leading to differences from the otherwise established strategy during the later blocks. If the anger movie is presented in later blocks, the strategy may have already been formed and thus harder to be influenced by states. Also, fatigue may have a dampening influence on state experience, leading to reduced anger induction in later experimental blocks. Additionally, this effect of the anger movie being presented first was biggest for fair offers, where it still led to a punishment reaction. This may be a hint that the anger induction was very powerful in this case and therefore participants kept on punishing prior unfair behavior even if the offer was now fair. We also found an influence of trait anger on punishment, with higher trait anger leading to more relative punishment, confirming findings already published by Rodrigues, Nagowski, et al. (2018). However, we lack to find the predicted interaction of trait anger with the state anger induction and therefore we found no evidence in this paradigm for trait activation by the anger induction (see e.g., Coan et al., 2006). This lack of interaction could be due to a too strong state induction (Rodrigues, Müller, et al. 2018). If an induced state is very strong, then everyone may react to this state similarly. An induction with lesser intensity, however, may lead to a more complex induction result, interacting with the relevant traits and only leading to an effect if the person is able to recognize and process the relevant information due to a preparedness in perception for the relevant trait related signals. Spider phobics, for example, can detect spider pictures better than healthy controls if they are presented for a short time interval (e.g., Pflugshaupt et al., 2005; Wieser, Pauli, Weyers, Alpers, & Mühlberger, 2009), but if a giant spider is put right before the participants, everyone may detect it. As the movie induction seems to be rather intense, it may supersede the trait interaction, although the trait has still an influence as predicted, leading to more punishment for higher trait anger. This leads to the already risen question whether punishment is really altruistic by its nature, as suggested by the term "altruistic punishment" (e.g., Fehr & Gächter, 2002). As Rodrigues, Nagowski, et al. (2018) already showed it is more likely to be driven by trait anger. In the present research, we add further evidence to this line of reasoning by showing that experimentally induced state anger also leads to higher punishment but not to higher compensation (see also Jordan et al., 2016; Nelissen & Zeelenberg, 2009; Pedersen et al., 2013; Seip et al., 2009, 2014). We conclude that the term "altruistic punishment" should not be used for this kind of behavior, and advocate the use of the term "costly punishment" (e.g., Brethel-Haurwitz et al., 2016; Henrich et al., 2006), as it does not include the trait or motivational concept altruism and focusses more on a pure description of the behavior.

Concerning the electrophysiological reaction to unfair offers, we found the hypothesized FRN effect to unfair offers for participants with higher trait anger that was not present in the midfrontal theta band activation. However, we did not find any difference in FRN to the anger induction movie sequence. Hence, persons with higher trait anger may react to unfair offers with a higher FRN because of a hypersensitive detection system for stimuli that may induce anger. As the unfair offer is such a stimulus, the FRN may be enhanced to this offer independent of the situation at hand. But as the state induced anger did not alter the FRN in any way, this leads to the conclusion that the mere state of anger is not sufficient to enhance electro-cortical reactions to unfair offers, while it is sufficient to alter the behavioral response. Thus, additional processes beyond the mere evaluation of the offer as fair or unfair contribute to the behavioral reaction, leading to higher punishment while showing no differences in the electro-cortical feedback response to the stimulus. This stresses also the point that despite the similar processing on electrophysiological basis, the subjective experience may differ very strongly, if emotional valence is brought into play. Angus et al. (2015), found similar results for the reward positivity (Rew-P), being independent of the anger induction in their study, too. Therefore, it is only partly surprising that the FRN was not influenced by the anger state manipulation, although it appeared to be rather strong for the behavioral results. But the feedback process underlying the FRN might as well only evaluate the fairness of the offer and as the participant is not directly affected, the influence of the emotion induction might not be as strong as if one would be the victim of the unfairness of other players (compare e.g., Weiß et al., 2019, for the ultimatum game). The third-party perspective might have prevented the influence of the anger induction on the FRN, as the participants only assess the fairness of the offer without suffering the consequences from unfair offers themselves. Hence, they might just react to the unfair offers in a more rational manner in contrast to what they would do if they were the receiver during the game.

For the single-trial midfrontal theta band activation as a predictor for behavior, we found a negative relation of midfrontal theta band activation with punishment behavior, which was not present for the single-trial FRN. This leads to two different conclusions: First, as midfrontal theta band activation has been linked to altruistic acts (e.g., Rodrigues et al., 2015) and punishment is not positively related to it, the notion that punishment is not linked to altruism is also given on the electrophysiological level. Besides the behavioral results, this electro-cortical finding strengthens the point of the punishment in third-party dictator games to be called "costly punishment" (e.g., Brethel-Haurwitz et al., 2016; Henrich et al., 2006; Rodrigues, Nagowski, et al., 2018) instead of "altruistic punishment" (Fehr & Gächter, 2002). The second important point arises from the comparison of the single-trial FRN with the single-trial theta band activation. In our previous work, we wanted to operationalize theta band activation as a proxy for FRN on trial-level (see also Cavanagh et al., 2012 for the relation of FRN with theta band activation), but the present results reveal a difference between the two indices. Following the findings of Cohen, Elger, and Ranagath (2007), we assume that theta band activation and the event-related components are two different approaches to identifying different aspects of perceptional, motivational, and even behavioral electrophysiological signal responses. If they are combined, they may lead to a deeper understanding and distinction of the different processes in a complex act like decision making. The FRN may hereby be more of evaluative nature (Holroyd & Coles, 2002), while midfrontal theta band activation may be more strongly linked to processes of cognitive control (e.g., Cavanagh et al., 2009; Cavanagh & Frank, 2014; Cohen, 2011; Phillips et al., 2014; Polanía et al., 2012). Hence the simultaneous use of FRN and theta band activation allowed for additional insights as we were able to show a separation of the two approaches to electrophysiological responses: The FRN was identified as the stimulus-related component of fairness perception concerning the offer of the dictator, while the theta band activation is negatively linked to the act of punishment. Hence, it is important to report both metrics of electro-cortical activation, in order to assess different aspects of responses and decision processes. The evaluative nature of the FRN is elucidated by being especially pronounced for unfair offers and being influenced by trait anger. But the midfrontal theta activation depicts decision processes that may need more cognitive control and cognitive effort (e.g., Cavanagh & Frank, 2014; Jensen & Tesche, 2002; Klimesch, 1999), like fair offers in a dictator game (see Rodrigues et al., 2015) or compensation instead of punishment in the third-party dictator game. As the punishment might follow some gut feeling of anger and lead to an impulsive decision (Strack & Deutsch, 2004), the act of not to punish needs to ignore this first feeling of anger and focus on not harming the unfair dictator (Mussel et al., 2013). A similar explanation might also be true for the previous findings of Rodrigues et al. (2015): While being a dictator, the gut feeling does not primarily lead to fair offers but rather to executing the power one has in this position, and accordingly the decision to not exploit the situation will need cognitive control. Hence, we interpret this difference in theta band activation for punishment versus. no punishment as a higher cognitive effort, being an electro-cortical correlate of an impulsive versus a more empathic or rational decision.

4.1 | Limitations

One great limitation of this study is the shortcoming concerning the induction of anger and its effect on behavior. Unfortunately, only if the anger movie was presented as the first movie, a behavioral effect could be detected. In further studies, one should try a more intense and personally relevant form of priming anger. One possibility would be to frustrate the participants with some fake information about topics like study tuitions or tax rise (e.g., Harmon-Jones, Sigelman, Bohlig, & Harmon-Jones, 2003). This more relevant and also money-related induction of anger may have a more intense impact on behavioral results, although a higher impact on the FRN is not likely (Angus et al., 2015).

Another limitation of this work is the concentration on negative feedback, the feedback-related negativity, while one could also take the perspective of the influence of reward and therefore focus on the reward positivity (Rew-P, Baker & Holroyd, 2011). However, as the task does not provide the participant with the money that is depicted in the offer and the offers are mostly not fair, the Rew-P framework might not be perfectly adequate in this case.

Yet another limitation is given by the sample size calculation. Having calculated the needed sample size only based on the main effect of punishment, a more complex effect like the interaction of trait and state anger might not reach sufficient power to be detected. In further studies, one should estimate the relevant interaction to calculate an appropriate sample size for this effect.

One further limitation is the duration of the experiment. As the participants experienced every movie condition, the experiment duration could have led to more strategic decisions in later blocks. But in order to achieve a randomized within design along with sufficient trials for an ERP analysis, we had to incorporate that many trials. Further studies could aim for a simpler design with only two conditions, leading to a less complex and shorter design. A between design would also be a possible solution, however, the great inter-individual variances in EEG would also lead to excessive sample sizes as well as very limited power. Hence, we would recommend a simpler, within participants design for further studies, possibly also including other emotion inductions that are more closely linked to the task at hand instead of the chosen more global manipulation. For example, a confederate might pose as dictator or receiver and briefly interact with the subject in a positive or negative manner in order to increase or decrease punishment or compensation, respectively. Also in future studies, a state manipulation of altruism should be used to further investigate the relation of state altruism with compensation and punishment.

Finally, one has to take into consideration that the effect of more punishment after the anger movies was partly driven by having more punishment in response to fair offers. Hence, participants might have used the punishment as a coping mechanism for their anger, independent of the fairness of the dictators' offers (see also Pillutla & Murninghan, 1996). However, these results would be in line with our previous point that the punishment is not driven by altruism but by anger, although the base rate of punishment is increased.

4.2 | Conclusion

In this study, we showed that punishment in a third-party dictator game is linked to state anger and it is not related to trait altruism. Also, we found on electro-cortical level that midfrontal theta band activation, which has been linked to altruistic acts previously, is negatively related to punishment. We interpret this finding as more cognitive control being needed to refrain from punishment, as punishment might be an impulsive decision. We further highlighted the importance of the distinction between the EEG-components FRN and midfrontal theta band activation, as they may offer different insight into the complex decision task of the third-party dictator game. The FRN was interpreted as an indicator of the perceived unfairness of an offer, while the single-trial midfrontal theta band activation was understood as an electro-cortical correlate of cognitive control and cognitive effort, leading to less impulsive behavior. Integrating these findings, it is vital to analyze event-related components along with the frequency response in tasks to get a more detailed view on complex behavior. Also, one should look at the different strengths and weaknesses of those two measures, especially if one is interested in analyzing single-trial EEG and EEG signals aggregated across conditions.

Furthermore, it is important to redefine the term "altruistic punishment" into terms like "costly punishment," in order to clearly distinguish the altruistic motive from a behavioral response that is mostly driven by anger. Support for this interpretation stems from the findings that neither the electro-cortical reactions nor the motivation for punishment is similar to an altruistic reaction in a third-party dictator game, if the participants are able to execute compensation and punishment behavior independently from each other.

CONFLICT OF INTEREST

No potential conflicts of interests exist.

ORCID

Johannes Rodrigues https://orcid.org/0000-0001-8471-0816
Mario Reutter https://orcid.org/0000-0002-5271-7594
Patrick Mussel https://orcid.org/0000-0001-5010-5677
Johannes Hewig https://orcid.org/0000-0002-8400-189X

REFERENCES

Ahlmann-Eltze, C. (2019). ggsignif: Significance Brackets for "ggplot2". Retrieved from https://cran.r-project.org/package=ggsignif

Andreoni, J. (1988). Why free ride? *Journal of Public Economics*, *37*(3), 291–304. https://doi.org/10.1016/0047-2727(88)90043-6

Angus, D. J., Kemkes, K., Schutter, D. J. L. G., & Harmon-Jones, E. (2015). Anger is associated with reward-related electrocortical activity: Evidence from the reward positivity. *Psychophysiology*, *52*(10), 1271–1280. https://doi.org/10.1111/psyp.12460

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

Boksem, M. A. S., & De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, *5*(1), 118–128. https://doi.org/10.1080/17470910903202666

Brethel-Haurwitz, K. M., Stoycos, S. A., Cardinale, E. M., Huebner, B., & Marsh, A. A. (2016). Is costly punishment altruistic? Exploring rejection of unfair offers in the Ultimatum Game in real-world altruists. *Scientific Reports*, *6*, 18974. https://doi.org/10.1038/srep18974

Carlo, G., Hausmann, A., Christiansen, S., & Randall, B. A. (2003). Sociocognitive and behavioral correlates of a measure of prosocial tendencies for adolescents. *The Journal of Early Adolescence*, *23*(1), 107–134. https://doi.org/10.1177/0272431602239132

Carlo, G., & Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of Youth and Adolescence*, *31*(1), 31–44. https://doi.org/10.1023/A:1014033032440

Carlo, G., Randall, B. A., Rotenberg, K. J., & Armenta, B. E. (2010). A friend in need is a friend indeed: Exploring the relations among trust beliefs, prosocial tendencies, and friendships. In *Interpersonal trust during childhood and adolescence* (pp. 270–294). Cambridge: Cambridge University Press.

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. B. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *Journal of Neuroscience*, *29*(1), 98–105. https://doi.org/10.1523/JNEUROSCI.4137-08.2009

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., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x

Coan, J. A., Allen, J. J. B., & McKnight, P. E. (2006). A capability model of individual differences in frontal EEG asymmetry. *Biological Psychology*, *72*(2), 198–207. https://doi.org/10.1016/J.BIOPSYCHO.2005.10.003

Cohen, M. X. (2011). Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *NeuroImage*, *55*(3), 1373–1383. https://doi.org/10.1016/J.NEUROIMAGE.2010.12.072

Cohen, M. X. (2014). *Analyzing neural time series data theory and practice*. (1st ed.). Cambridge, Massachusetts, London, England: The MIT Press.

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, *35*(2), 968–978. https://doi.org/10.1016/J.NEUROIMAGE.2006.11.056

Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *Journal of Neuroscience*, *27*(2), 371–378. https://doi.org/10.1523/JNEUROSCI.4421-06.2007

Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I. *Journal of Personality and Social Psychology*, *58*(2), 330–341. https://doi.org/10.1037/0022-3514.58.2.330

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of 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

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, *101*(1), 91–119. https://doi.org/10.1037/0033-2909.101.1.91

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370. https://doi.org/10.1177/1754073911410740

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140. https://doi.org/10.1038/415137a

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

Güth, W. (1995). On ultimatum bargaining experiments—A personal review. *Journal of Economic Behavior & Organization*, *27*(3), 329–344. https://doi.org/10.1016/0167-2681(94)00071-L

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*(2), 148–154. https://doi.org/10.1016/j.biopsycho.2005.04.001

Harmon-Jones, E., & Allen, J. J. B. (1998). Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology*, *74*(5), 1310–1316. https://doi.org/10.1037//0022-3514.74.5.1310

Harmon-Jones, E., Sigelman, J., Bohlig, A., & Harmon-Jones, C. (2003). Anger, coping, and frontal cortical activity: The effect of coping potential on anger-induced left frontal activity. *Cognition & Emotion*, *17*(1), 1–24. https://doi.org/10.1080/02699930302278

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., … Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767–1770. https://doi.org/10.1126/science.1127333

Henry, L., Wickham, H., & Chang, W. (2018). ggstance: Horizontal "ggplot2" components. Retrieved from https://cran.r-project.org/package=ggstance

Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., & Bartussek, D. (2005). A revised film set for the induction of basic emotions. *Cognition & Emotion*, *19*(7), 1095–1109. https://doi.org/10.1080/02699930541000084

Hewig, J., Hagemann, D., Seifert, J., Naumann, E., & Bartussek, D. (2004). On the selective relation of frontal cortical asymmetry and anger-out versus anger-control. *Journal of Personality and Social Psychology*, *87*(6), 926–939. https://doi.org/10.1037/0022-3514.87.6.926

Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltner, W. H. R. (2011). Why humans deviate from rational choice. *Psychophysiology*, *48*(4), 507–514. https://doi.org/10.1111/j.1469-8986.2010.01081.x

Hewig, J., Trippe, R., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltner, W. H. R. (2007). Decision-making in blackjack: An electrophysiological analysis. *Cerebral Cortex*, *17*(4), 865–877. https://doi.org/10.1093/cercor/bhk040

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., 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

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*(8), 1395–1399. https://doi.org/10.1046/j.1460-9568.2002.01975.x

Jordan, J., McAuliffe, K., & Rand, D. (2016). The effects of endowment size and strategy method on third party punishment. *Experimental Economics*, *19*(4), 741–763. https://doi.org/10.1007/s10683-015-9466-8

Kayser, J., & Tenke, C. E. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clinical Neurophysiology*, *117*(2), 348–368. https://doi.org/10.1016/j.clinph.2005.08.034

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2–3), 169–195. https://doi.org/10.1016/S0165-0173(98)00056-3

Kujawa, A., Smith, E., Luhmann, C., & Hajcak, G. (2013). The feedback negativity reflects favorable compared to nonfavorable outcomes based on global, not local, alternatives. *Psychophysiology*, *50*(2), 134–138. https://doi.org/10.1111/psyp.12002

Leiner, D. J. (2016). SoSci Survey (Version 2.6.00) [Computer software]. Retrieved from https://www.soscisurvey.de

Leliveld, M. C., Dijk, E., & Beest, I. (2012). Punishing and compensating others at your own expense: The role of empathic concern on reactions to distributive injustice. *European Journal of Social Psychology*, *42*(2), 135–140. https://doi.org/10.1002/ejsp.872

Lenth, R. V. (2016). Least-squares means: The R package {lsmeans}. *Journal of Statistical Software*, *69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Lotz, S., Baumert, A., Schlösser, T., Gresser, F., & Fetchenhauer, D. (2011). Individual differences in third-party interventions: how justice sensitivity shapes altruistic punishment. *Negotiation and Conflict Management Research*, *4*(4), 297–313. https://doi.org/10.1111/j.1750-4716.2011.00084.x

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210. https://doi.org/10.1016/j.tics.2004.03.008

Mathur, V. A., Harada, T., Lipke, T., & Chiao, J. Y. (2010). Neural basis of extraordinary empathy and altruistic motivation. *NeuroImage*, *51*(4), 1468–1475. https://doi.org/10.1016/j.neuroimage.2010.03.025

Miltner, W. H. R., 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

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*(2), 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x

Mothes, H., Enge, S., & Strobel, A. (2016). The interplay between feedback-related negativity and individual differences in altruistic punishment: An EEG study. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(2), 276–288. https://doi.org/10.3758/s13415-015-0388-x

Mussel, P., Göritz, A. S., & Hewig, J. (2013). Which choice is the rational one? An investigation of need for cognition in the ultimatum game. *Journal of Research in Personality*, *47*(5), 588–591. https://doi.org/10.1016/J.JRP.2013.05.007

Mussel, P., Reiter, A. M. F., Osinsky, R., & Hewig, J. (2015). State- and trait-greed, its impact on risky decision-making and underlying neural mechanisms. *Social Neuroscience*, *10*(2), 126–134. https://doi.org/10.1080/17470919.2014.965340

Mussel, P., Ulrich, N., Allen, J. J. B., Osinsky, R., & Hewig, J. (2016). Patterns of theta oscillation reflect the neural basis of individual differences in epistemic motivation. *Scientific Reports*, *6*(1), 29245. https://doi.org/10.1038/srep29245

Nelissen, R. M. A., & Zeelenberg, M. (2009). Moral emotions as determinants of third-party punishment: Anger, guilt, and the functions of altruistic sanctions. *Judgment and Decision Making*, *4*(7), 543–553. Retrieved from http://journal.sjdm.org

Pedersen, E. J., Kurzban, R., & McCullough, M. E. (2013). Do humans really punish altruistically? A closer look. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1758), 20122723–20122723. https://doi.org/10.1098/rspb.2012.2723

Pflugshaupt, T., Mosimann, U. P., von Wartburg, R., Schmitt, W., Nyffeler, T., & Müri, R. M. (2005). Hypervigilance–avoidance pattern in spider phobia. *Journal of Anxiety Disorders*, *19*(1), 105–116. https://doi.org/10.1016/J.JANXDIS.2003.12.002

Phillips, J. M., Vinck, M., Everling, S., & Womelsdorf, T. (2014). A long-range fronto-parietal 5- to 10-Hz network predicts "top-down" controlled guidance in a task-switch paradigm. *Cerebral Cortex*, *24*(8), 1996–2008. https://doi.org/10.1093/cercor/bht050

Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes*, *68*(3), 208–224. https://doi.org/10.1006/OBHD.1996.0100

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2018). {nlme}: Linear and nonlinear mixed effects models. Retrieved from https://cran.r-project.org/package=nlme

Polanía, R., Nitsche, M. A., Korman, C., Batsikadze, G., & Paulus, W. (2012). The importance of timing in segregated theta phase-coupling for cognitive performance. *Current Biology*, *22*(14), 1314–1318. https://doi.org/10.1016/J.CUB.2012.05.021

Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civai, C., Sartori, G., & Rumiati, R. (2008). Mentalizing in economic decision-making. *Behavioural Brain Research*, *190*(2), 218–223. https://doi.org/10.1016/J.BBR.2008.03.003

Polezzi, D., Lotto, L., Daum, I., Sartori, G., & Rumiati, R. (2008). Predicting outcomes of decisions in the brain. *Behavioural Brain Research*, *187*(1), 116–122. https://doi.org/10.1016/J.BBR.2007.09.001

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

Qu, C., Wang, Y., & Huang, Y. (2013). Social exclusion modulates fairness consideration in the ultimatum game: An ERP study. *Frontiers in Human Neuroscience*, *7*, 505. https://doi.org/10.3389/fnhum.2013.00505

R Core Team. (2019). R. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rodrigues, J., Müller, M., Mühlberger, A., & Hewig, J. (2018). Mind the movement: Frontal asymmetry stands for behavioral motivation, bilateral frontal activation for behavior. *Psychophysiology*, *55*(1), e12908. https://doi.org/10.1111/psyp.12908

Rodrigues, J., Nagowski, N., Mussel, P., & Hewig, J. (2018). Altruistic punishment is connected to trait anger, not trait altruism, if compensation is available. *Heliyon*, *4*(11), e00962. https://doi.org/10.1016/J.HELIYON.2018.E00962

Rodrigues, J., Ulrich, N., & Hewig, J. (2015). A neural signature of fairness in altruism: A game of theta? *Social Neuroscience*, *10*(2), 192–205. https://doi.org/10.1080/17470919.2014.977401

Rodrigues, J., Ulrich, N., Mussel, P., Carlo, G., & Hewig, J. (2017). Measuring prosocial tendencies in Germany: Sources of validity and reliablity of the revised prosocial tendency measure. *Frontiers in Psychology*, *8*, 2119. https://doi.org/10.3389/fpsyg.2017.02119

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, *300*(5626), 1755–1758. https://doi.org/10.1126/science.1082976

Schellberg, D., Besthorn, C., Klos, T., & Gasser, T. (1990). EEG power and coherence while male adults watch emotional video films. *International Journal of Psychophysiology*, *9*(3), 279–291. https://doi.org/10.1016/0167-8760(90)90060-Q

Schwenkmezger, P., & Hodapp, V. (1991). A questionnaire for assessing anger and expression of anger. *Zeitschrift Fur Klinische Psychologie, Psychopathologie Und Psychotherapie*, *39*(1), 63–68. Retrieved from http://psycnet.apa.org/record/1991-77110-001

Seip, E. C., Van Dijk, W. W., & Rotteveel, M. (2009). On hotheads and dirty harries: The primacy of anger in altruistic punishment. *Annals of the New York Academy of Sciences*, *1167*(1), 190–196. https://doi.org/10.1111/j.1749-6632.2009.04503.x

Seip, E. C., Van Dijk, W. W., & Rotteveel, M. (2014). Anger motivates costly punishment of unfair behavior. *Motivation and Emotion*, *38*(4), 578–588. https://doi.org/10.1007/s11031-014-9395-4

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2019).afex: Analysis of factorial experiments. Retrieved from https://cran.r-project.org/package=afex

Spielberger, C. D. (1988). *Manual for the state train anger expression inventory*. Odessa, FL (P.O. Box 998 Odessa 33556): Psychological Assessment Resources.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1

Sun, L., Tan, P., Cheng, Y., Chen, J., & Qu, C. (2015). The effect of altruistic tendency on fairness in third-party punishment. *Frontiers in Psychology*, *6*, 820. https://doi.org/10.3389/fpsyg.2015.00820

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Harlow, Essex: Pearson.

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*(3), 500–517. https://doi.org/10.1037/0021-9010.88.3.500

Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality*, *34*(4), 397–423. https://doi.org/10.1006/jrpe.2000.2292

Weiland, S., Hewig, J., Hecht, H., Mussel, P., & Miltner, W. H. R. (2012). Neural correlates of fair behavior in interpersonal bargaining. *Social Neuroscience*, *7*(5), 537–551. https://doi.org/10.1080/17470919.2012.674056

Weiß, M., Gutzeit, J., Rodrigues, J., Mussel, P., & Hewig, J. (2019). Do emojis influence social interactions? Neural and behavioral responses to affective emojis in bargaining situations. *Psychophysiology*, *56*, e13321. https://doi.org/10.1111/psyp.13321

West, S. G., Ryu, E., Kwok, O.-M., & Cham, H. (2011). Multilevel modeling: Current and future applications in personality research. *Journal of Personality*, *79*(1), 2–50. https://doi.org/10.1111/j.1467-6494.2010.00681.x

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, *26*(4), 557–580.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag. Retrieved from http://ggplot2.org

Wickham, H., & Bryan, J. (2019). readxl: Read excel files. Retrieved from https://cran.r-project.org/package=readxl

Wieser, M. J., Pauli, P., Weyers, P., Alpers, G. W., & Mühlberger, A. (2009). Fear of negative evaluation and the hypervigilance-avoidance hypothesis: An eye-tracking study. *Journal of Neural Transmission*, *116*(6), 717–723. https://doi.org/10.1007/s00702-008-0101-0

Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behavioral and Brain Functions*, *7*(1), 30. https://doi.org/10.1186/1744-9081-7-30

Wu, Y., Leliveld, M. C., & Zhou, X. (2011). Social distance modulates recipient's fairness consideration in the dictator game: An ERP study. *Biological Psychology*, *88*(2–3), 253–262. https://doi.org/10.1016/j.biopsycho.2011.08.009

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

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

TABLE S1 Greenhouse-Geisser corrections for ANCOVAs
TABLE S2 Marginal effects for different dependent and independent variables for ANCOVAs
TABLE S3 Model fit criteria for the different single trial multilevel models for relative compensation and relative punishment, including single trial FRN as predictor. The bold models were selected for further analyses
TABLE S4 Model fit criteria for the different single trial multilevel models for relative compensation and relative punishment, including midfrontal theta band activity as predictor. The bold models were selected for further analyses
TABLE S5 Marginal effects for different dependent variables for the single trial FRN multi-level model
FIGURE S1 Subjective ratings dependent on the different movie sequences for perceived anger, negative emotions and positive emotions. Error-bars represent mean *SE*

How to cite this article: Rodrigues J, Liesner M, Reutter M, Mussel P, Hewig J. It's costly punishment, not altruistic: Low midfrontal theta and state anger predict punishment. *Psychophysiology*. 2020;57:e13557. https://doi.org/10.1111/psyp.13557