Action Monitoring in One-Dimensional Force Production

Inauguraldissertation

zur

Erlangung des Doktorgrades

der Humanwissenschaftlichen Fakultät

der Universität zu Köln

nach der Promotionsordnung vom 18.12.2018

vorgelegt von

Yohana Diah Laksmi Siswandari

aus Jakarta (Indonesien)

Abgabe: Juli 2019

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln im November 2019 angenommen.

Gutachter*innen: Prof. Dr. Jutta Stahl
Prof. Dr. Stefan Bode (University of Melbourne)

Tag der mündlichen Prüfung: 29. November 2019

Contents

Acknowledgements

I would like to thank everybody who have helped me, not only during the writing of this thesis, but also those who have stood beside me over the course of this PhD journey. First of all, I would like to wholeheartedly thank my supervisor, Jutta Stahl, who has shared her expertise and guided me throughout my PhD journey, and for providing countless support and help during my study. I am forever grateful for having given an opportunity to be part of this research group. I would also like to express my gratitude to Stefan Bode, who has shared his expertise regarding the multivariate approach used in this study, and has provided invaluably constructive criticism and advice for my work.

Without the financial support of Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan / LPDP RI) and the Indonesian Ministry of Finance, which offered me a full four-years doctorate scholarship, this work would not have been possible. I am forever grateful for this financial support, and I hope that this doctoral thesis provides a valuable knowledge that will be useful to the development of Cognitive Neuroscience research in Indonesia.

Next, I would also like to thank my colleagues for helping me at one or more parts of my PhD journey. First, I am grateful to Sebastian Dummel, for sharing his knowledge related to Psychological Refractory Period and gave me ideas to improve my second study. I would also like to thank André Mattes, for helping me with the models' construction for study 2. I am also deeply grateful to my other colleagues: Kilian Kummer, Manuela Hundrieser, Markus Mück, Ulrike Thesing, Wilfried Follman, for the numerous discussions on both studies.

Last but not least, I would like to express my deepest gratitude to my family and friends. My parents and my grandparents, who always believed in me and supported me to get through the difficult times in my PhD journey, and my partner, Benni, who provides me with unconditional love, and for always cheering me up during difficult times. I would also like to thank my best friend, Niessa, for giving me daily dose of emotional support and for always making jokes that never failed to uplift me during hard situations. I cannot express enough gratitude for having such a great friend.

Abstract

An important mechanism that is essential to avoid error and is possessed by human beings is the ability to monitor one's own cognitive functioning, often referred to as 'action monitoring'. Force production is one example from numerous real life situations which requires precise and continuous action monitoring, in which an irreversible false decision during its execution could lead to highly dangerous situations. Despite its importance, most of the studies in the area of action monitoring investigated tasks with discrete response parameters (i.e., *left* or *right* hand responses during flanker task), leaving action monitoring during continuous tasks like force production insufficiently investigated. The general aim of the present research was to learn more about the response dynamics during a simple one-dimensional force production task so that the monitoring processes involved in the brain during the task execution could be better understood. Therefore, two studies were conducted. The first study (N = 48) investigated how force execution of a simple one-dimensional force production task unfolds. The results from this study indicated that magnitude and timing of the force pulse (i.e., the response force parameters) were defined by the motor program even before response execution. As a force pulse is a ballistic process, an early definition of the response parameters seems to be an efficient monitoring strategy, as it allows for a fast force production and could provide information for error detection process. However, the process of determining these response parameters seemed to precede the process of determining the *correctness* of the response itself. The second study (N = 40) incorporated a modified *Psychological Refractory Period* (PRP) paradigm using two consecutive simple force production tasks, to investigate if force monitoring process during these two subsequent force tasks was modulated by different Response-Stimulus Intervals. This study showed successful replication of a PRP effect in an adapted force paradigm. On a behavioral level, PRP effect seemed to modulate not only Response Time, but also the response force parameters such as Peak Force and Time-to-Peak Force. Furthermore, PRP effect also seemed to modulate the neural correlates of force monitoring. Finally, three alternative models were postulated to elucidate how the different Response-Stimulus Interval affected the information processing stages of two subsequent simple one-dimensional force tasks. Taken together, findings from the second study could serve as an evidence that PRP effect is present in subsequent simple force production tasks, without the necessity of motor overlap between the first and the second task.

I. Introduction and Theoretical Background

1.1 Errors vs. Efficient Performance Monitoring

The human body is a complex system comprising billions of cells which interact with one another, creating subsystems that enable an individual to perform different tasks in everyday life. To respond to environmental contingencies, human requires not only a system that generates a response in a certain way, but also a sufficient monitoring system to assure that an action is performed properly. When the monitoring system fails to perform its function properly or is somehow disrupted, an ‘error’ occurs. An ‘error’ is defined as the failure of a planned sequence of mental or physical activities to achieve its intended outcome when these failures cannot be attributed to chance (Reason, 1990). Generally, errors depend on two types of failure: (1) actions do not go as intended; (2) the intended actions are not the correct one. What differentiates these two types of errors is, for the first case, the desired outcome may or may not be achieved, while for the second case, the desired outcome cannot be achieved. Note that this study was focused on the first definition of error, thus the second type of error is irrelevant.

Human being is a big contributor to errors in a working system. In term of understanding how human ‘contributes’ to an error, it is important to separate *active* and *latent* errors. *Active error* happens at the level of a frontline operator, whose effects can be felt almost immediately, while *latent errors* includes things such as poor design, incorrect installation, faulty maintenance, incorrect management decisions, and poorly structured organizations. A good example of both errors can be illustrated by a plane crash. Here, the crash is a consequence of a committed error. This error can be categorized as an *active* or a *latent* error, depending on how the situation unfolds. In this case, an *active error* is committed when the pilot crashed the plane unintentionally, for instance, due to a lack of attention or drowsiness. On the other hand, a *latent error* is committed when a previously unknown design malfunction (i.e., a faulty sensor that erroneously reported that the airplane was stalling during *take-off*) caused the system to automatically maneuver the plane strangely in a way that the pilot (i.e., the frontline operator) could not control, which resulted in a plane crash. Often, only ensuring that an *active error* is not committed does not guarantee for a safe system. *Latent errors* are more likely to be ‘hidden’ in a system design of a routine process, thus posing a big threat to safety in a complex system, since they are often unrecognized and can result in multiple active errors. That being said, it is actually terrifying to think that even an action is done properly, a danger in numerous situations in life is sometimes unavoidable. A widely known real-life situation that exhibited how *latent error* occurred and how significant the consequences were, is the Chernobyl nuclear disaster, which was a result of a design flaw in the reactor system combined with human errors. Thus, a thorough understanding of a specific system is necessary to avoid unforeseeable error(s) that can lead to multiple *active errors* and potentially dangerous situation.

An important mechanism that is essential to avoid errors and is possessed by human beings is the ability to monitor one's own cognitive functioning (from now on, referred to as *monitoring*). In order to successfully conduct a motor action (for example, a simple motor action like moving from one point to another), human brain needs a mental representation of how this action could be done, using the available information (i.e., in this case: sensory information, memory, possible routes, means to get to the other spot, possible obstacles, *et cetera*). To create this mental representation, an adequate action monitoring system plays a significant role in detecting movement error and planning an appropriate reaction (i.e. immediately modify inappropriate movements, immediately stop before bumping into some obstacles). When access to information is limited for evaluation, or when organizing information becomes a difficulty, a monitoring process can be impaired (Chan et al., 2015) and therefore the probability of an error commission is increased.

Numerous real life situations require very precise and continuous action monitoring, when the simplest mistakes could possibly lead to highly dangerous situations. One of many situations where an accurate monitoring system is essential to avoid dangerous situations is while producing force. Note that force production itself is regarded as a 'continuous task', since it is a time-related activity and therefore requires continuous motor control during its execution (i.e., determining the response force parameters such as peak force and time-to-peak force) until the Peak Force is reached. Consequently, errors in force production are not exclusively errors in the process of selecting the 'appropriate response', which is usually the case in discrete choice tasks such as flanker or go/no-go tasks. In a force production task, there are two kinds of errors: (1) errors in response selection (e.g., while deciding to produce a high or low response force); (2) errors during the 'program' execution (e.g., failure to produce the exact force units planned during a specific time interval, which leads to under/over-producing the response force). A rather ‘spectacular’ example that illustrates the necessity of continuous monitoring process during force execution would be, when a pilot of a fighter aircraft needs to apply a particular amount of force on the rudder pedal in order to maintain a straight and level flight while moving in a high speed. Another occasion when a precise and continuous force monitoring is highly necessary would be during a surgical operation, where the surgeon (or a machine controlled by surgeon) needs to apply an accurate amount of force when cutting a patient’s flesh and not damage the nerves or organs in a close proximity. However, even in less spectacular situations, like chopping vegetables or peeling fruits with a knife, an efficient force monitoring system is essential.

1.2 Performance Monitoring in a Simple One-dimensional Force production

As previously described, accurate force production requires a good synergy of not only cognitive and motor processes, but also efficient monitoring system. In order to obtain basic understanding regarding the nature of force production itself, one should start from the simplest task. This study started with an investigation of performance monitoring during a simple one-dimensional force production task (i.e., achieved through a simple key press along the vertical axis, enabling upward or downward finger movement). In this section, a general overview of various means of investigation used in this study were outlined.

1.2.1 Univariate Approach: The Use of Event-Related Potentials

Electroencephalography (EEG) allows for measuring the brain electrical fields via metal electrodes placed on the human’s head. The electrical fields were resulted from electro-chemical signals passing from one neuron to another. These so called electrical fields would be powerful enough to be measurable using noninvasive instruments (i.e., EEG devices) when a large amount (billions) of these very little signals are passed simultaneously in neural populations which are aligned geometrically, and are spatially extended (Cohen, 2017). So, what exactly ignites this electrical field? Basically, when a neuron receives new information, an electrical signal is transmitted. This electrical signal triggers the release of chemicals known as neurotransmitter at special locations (known as *synapses*). As previously mentioned, a noninvasive instrument such as an EEG device can be used to measure the electrical signaling between the neurons via metal electrodes, usually used with a conductive gel to ensure a better contact between the metal electrodes and the scalps. Whenever a group of neurons are activated, local current flows (which vary in time) are produced (this flows consist of mainly Na+, Ca++, K+, and CL-). Every neuron creates a tiny electrical dipole between the soma (neuron's body) and the apical dendrites (the neural branches), and this will cause the current flow to be pumped through the membrane channels of the neurons, in the direction governed by the membrane's potential (Teplan, 2002). This process leads to excitatory or inhibitory postsynaptic potentials that add together in the cortex and extend to the surface of the scalp, resulting in a measurable voltage at the scalp (Luck, 2005). This voltage is what measured by EEG devices.

Event-related potentials (ERPs) are described as very small voltages in the brain in response to specific events such as stimuli (Blackwood & Muir, 1990). ERP signals are actually EEG changes which are time-locked to sensory, motor, or cognitive events, and provide a safe non-invasive approach to investigate neural correlates of certain mental processes. They allow for observation of cognitive operations that happen from before the sensory information is delivered to the peripheral nervous system, until after a behavioural response is made. According to Peterson et al. (1995), ERPs reflect the summed activity of postsynaptic potentials produced when a large amount (thousands or even millions) of cortical pyramidal neurons fire in synchrony while information is being processed. In human beings, ERPs can be divided into 2 categories. The ERP signals peaking during the first 100 ms after a stimulus presentation are called 'sensory' or 'exogenous', since these signals are largely dependent on the physical properties of the stimulus. On the other hand, ERPs which are generated later reflect the moment when the stimulus is being evaluated – and are known as 'cognitive' or 'endogenous' ERPs since they reflect information processing. To extract ERPs from a vast amount of EEG data, the time points when a stimulus occurs and when a response is initiated have to be marked (as stimulus and response triggers). Using these triggers, a large amount of specific time-windows can be extracted from the continuous EEG data. These time windows are known as 'epochs', and are usually time-locked to one of these triggers (depending on the needs). The length of 1 epoch is usually pre-defined (e.g., -300 ms before until 300 ms after a response is triggered). This whole procedure of extracting epochs from a single continuous EEG data is called 'EEG epoching'. The obtained epochs are then averaged. This averaging is important, since it allows for filtering out all activities in the brain signal that are not related to the event of interest. To average out these unrelated signals, of course, an adequate amount of trials is required.

Averaged ERP waveforms consists positive and negative deflections. A significant positive or negative deflection is termed as ‘component’. Usually, a component is indicated by its polarity and the approximate time points when it happens. For instance, a negative deflection that occurs approximately 100 ms after stimulus presentation is called “N100” (the letter ‘N’ indicates polarity, and the number ‘100’ indicates the time point when this deflection is observable relative to the onset of the event). Until recently, various ERP components were discovered and these component are known to be associated with certain processes. In force production paradigms, several components have been identified as indicators of different aspects of error processing as well as error-specific variations. In below sections, the roles of these components in force monitoring are discussed.

1.2.1.1 Response-locked Event-Related Potentials

The first and the most used ERP component in performance monitoring studies is the frontocentral *error*-(*related*) *negativity* (Ne/ERN), a sharp negative deflection peaking between 0 to 180 ms after error response onset (see Falkenstein et al., 1990). This component was said to be originated from the *anterior cingulate cortex* (ACC), which is located in the medial frontal cortex (see Dehaene, Posner, & Tucker, 1994; Ullsperger & von Cramon, 2001; Yeung et al., 2004). Ne/ERN reflects early error processing mechanisms. A component with similar characteristics, the *correct response negativity* (CRN), occurs after correct responses and has been found to index response uncertainty (e.g., Vidal, et al., 2000).

The *error positivity* (Pe), which is an indicator of an error detection mechanism related to error awareness, peaks around 300 ms after an erroneous response (Nieuwenhuis et al., 2001; Steinhauser & Yeung, 2010). Although Pe has been neglected in recent force monitoring studies, it constitutes an interesting behavioural adaptation, error awareness and error evidence accumulation in general error-monitoring studies. For example, Nieuwenhuis et al. (2001) pointed out that the amplitude of Pe (but not the Ne/ERN) covaried to the degree of error awareness. This is supported by another study from Leuthold & Sommer (1998), which also provided evidence that Pe amplitude correlated with the salience of information that induced errors. Nieuwenhuis et al. (2001) also reported that Pe correlated with behavioural adaptation after a committed error (i.e. the *post-error slowing*). What makes it particularly interesting to incorporate Pe into force monitoring is the fact that, unlike Ne/ERN that usually peaks before the peak force of a brief force pulse is reached, Pe usually peaks around 300 ms after the response onset. This is the point where the peak force of a brief force pulse is (often) already reached – making it a potentially interesting neural indicator for force-related error processing mechanisms.

1.2.1.2 Feedback-locked Event-Related Potentials

Ensuring that a force is produced accurately is a complex process, therefore most of the time human (aside from motor-experts) needs not only internal but also external monitoring, which is provided by external feedback following a force response. Feedback related negativity (FRN) is a widely known ERP component which was expected to follow the feedback in each trial, as this component reflects external error processing. FRN can be observed around 250 ms after an error feedback onset. According to the *reinforcement learning theory*, this component indicates prediction errors, such as if an outcome was worse than the expectation (Holroyd & Coles, 2002). FRN could also reflect error-related information that were not reflected in the Ne/ERN (i.e., *first-indicator hypothesis*, see Holroyd & Coles, 2002).

FRN is often followed by the Feedback P3, which was defined as a positive deflection around 300 ms after the feedback onset. The literature suggested that Feedback P3 is affected by factors that also influence FRN (Sailer et al., 2010), in particular the conformity to expectations. In reward-related processes, an increased Feedback P3 followed a larger magnitude of win or loss feedback (e.g., Yeung and Sanfey, 2004; Yeung et al., 2004, Hajcak et al., 2005). However, in term of directions, the literature also provided conflicting findings. For instance, Bellebaum and Daum (2008) as well as Hajcak et al. (2005) found a larger P3 amplitude only for the positive feedback, while Frank et al. (2005) observed an increase P3 amplitude for negative feedback. Furthermore, Holroyd et al. (2003) observed larger P3 amplitudes when reward was unexpected (i.e., larger than expected).

1.2.2 Multivariate Approach: The Use of Multivariate Pattern Analyses

Aside from the classical ERP components, the use of *machine learning-based* techniques has been increasingly popular nowadays in neuroscience research. Machine learning itself is an implementation of algorithms and statistical models that computer systems use to perform a specific task effectively without using *a priori* knowledge or explicit instructions, usually relying on patterns and inference. A mathematical model (based on a 'training set' of a specific sample data) is created through the use of specific algorithms, in order to create a decision boundary that can be used to categorize data with similar characteristics/ patterns to different groups. Note that this decision boundary does not have to be linear. Machine learning is often seen as a subset of artificial intelligence, and is currently applied in a wide range of problems. Email filtering is one example of machine learning application in everyday life.

One implementation of these machine learning algorithms in a form of chronometric, multivariate approach (multivariate pattern analysis / MVPA; see Bode et al., 2012; Bode & Stahl 2004) was used in the current study. To analyze neural data, MVPA classifiers are trained to predict information contained in the cognitive processes – directly from local spatial activation patterns. Despite the fact that MVPA has been used since decades ago, this method is currently experiencing vast development in term of its techniques. Recent studies have incorporated MVPA to analyze functional magnetic resonance imaging (fMRI) and EEG data (e.g., Haynes & Rees, 2006; Tong & Pratte, 2012; Bode & Stahl, 2014). Note that in comparison to fMRI, the use of MVPA on EEG data provides a far better temporal resolution (in the range of milliseconds). Utilizing the multivariate nature of EEG signals, different techniques of MVPA were used to predict stimulus types, decision outcomes, decision models' parameters, as well as errors – all directly from the obtained brain activity patterns (e.g., Bode & Stahl, 2014; Bode et al., 2012; Bode et al., 2014; Philiastides et al., 2006; van Vugt et al., 2012;). The advantage of this approach becomes particularly apparent when no variation in ERP components was found to be linked to the specific cognitive process of interest. In other words, the use of MVPA does not require existing knowledge of the exact timing and location of a cognitive process, allowing a more explorative search for information related to the event of interest.

1.3 Performance Monitoring in Two Successive Simple One-dimensional Force Productions: Challenges of the Psychological Refractory Period

Aside from investigating how a simple force production involving a single button press unfolds, it is of the author's interest to see what happened when a simple one-dimensional force was produced successively in a very short time. In the context of 'two successive tasks', the current work will rely on a widely known dual task paradigm –the so called *Psychological refractory period* (PRP), which is described as a period of time during which the response to a second stimulus is significantly slower because the first stimulus is still being processed. A PRP paradigm, as described by Pashler (1994), required two successive speeded responses to be made, while the time interval between the two stimulus onsets, known as the *Stimulus Onset Asynchrony* (SOA) is varied. A PRP effect is induced when the SOA between the two stimuli is reduced, resulting in a slowing of the second response.

So far, there are two popular theories that explain this effect. The first theory, the *Central Capacity Sharing* theory, postulates that dual-task interference is caused by a capacity limited process that allocates capacity in a graded fashion (Joliccoeur, 2002). To put it simply, the performance deficits during a reduced SOA emerge since there is a capacity limit of processes that can be run at the same time in the central processing mechanisms. These processes are, for example, memory encoding and retrieval, response selection, and other cognitive processes, which can originate from one or more tasks at the same time and happen before a motor response. A performance deficit (i.e., slowing in reaction time for the second task because of this shared cognitive capacity with first task) occurs when the central processing capacity reaches its maximum limit to process the cognitive processes required to execute the task(s) in a short period of time. What is noteworthy, this model considers that all stages of the two tasks, including the *Central Processing* stage such as selection of response, can proceed in parallel. This theory was supported by several empirical studies which demonstrated that the human information processing system is, in a number of different ways, sharply capacity-limited. For instance, a study by Miller (1956) suggested that one can only hold a limited amount of information in the short-term memory. However, other aspects of the processing system, for instance, the early stages of visual perception (Barbur et al., 1993), seem to show very large capacity.

For these aspects (of the processing system) with large capacity, capacity limitations could be the factor that induce bottlenecks in the flow of information processing. This theory is illustrated in Figure 1.

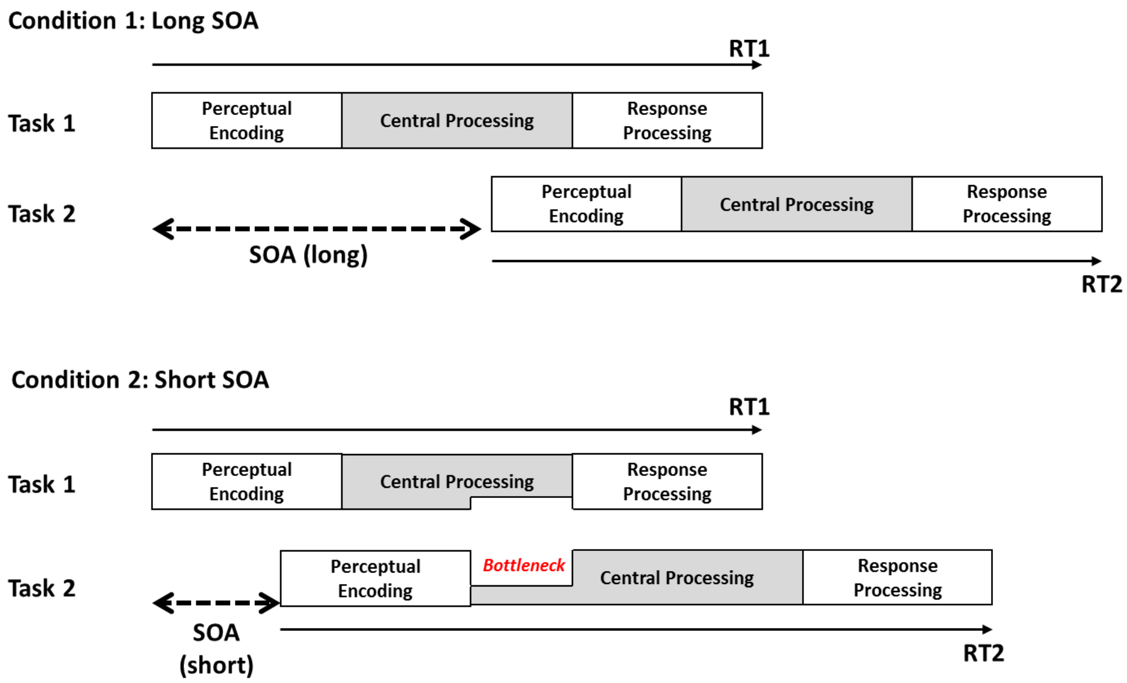


Figure 1. An Illustration of the Central Capacity Sharing Theory. Adapted from Figure 1B in Joliccoeur (2002, p. 3)

Although some researchers (e.g., Heil, et al., 1999; Jolliccoeur, 2002; Meyer & Kieras, 1997; Tombu & Jolliccoeur, 2005) believe that people can parallel multitask, laboratory studies (e.g., Pashler, 1992; Pashler, 1994) suggested otherwise. Researchers nowadays (e.g., Lien & Schweickert, 2003; Ruthruff, et al., 2003; Sigman & Dehaene, 2006; Jentzsch et al, 2007; Ulrich & Miller, 2008) seem to prefer the second theory, the *Central Bottleneck Theory* (see Welford, 1952) to explain PRP effect. This theory suggested that central mental processing can only process one task at a time. As a result, when a person needs to do the *Central Processing* stage for two tasks within a short period of time, the *Central Processing* stage of Task 2 is postponed, resulting in a slower reaction time for Task 2 (see Figure 2). It is important to note that, the only stage where an overlap is deemed 'impossible' to happen according to this model is the *Central Processing* stage. Thus, an overlap in other stages is possible (e.g., *Perceptual Encoding* for Task 1 and Task 2 could happen at the same time).

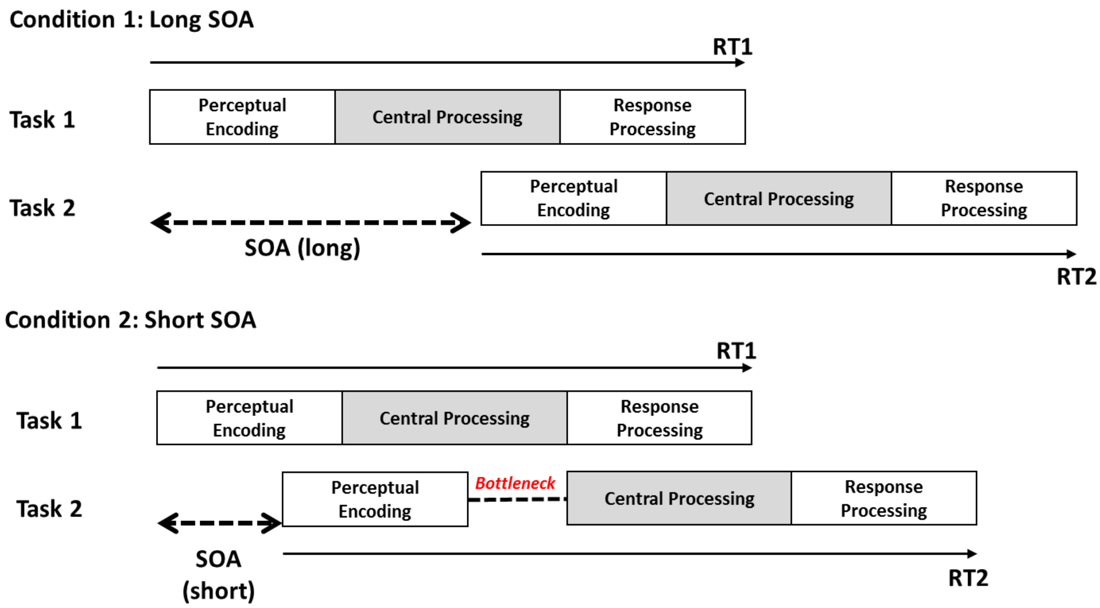


Figure 2. An Illustration of the Central Bottleneck Theory. Adapted from Figure 2 in Pashler (1994, p. 5)

In Figure 2, a cognitive bottleneck was illustrated to happen during the short SOA condition, when *Perceptual Encoding* of Task 2 finished before *Central Processing* of Task 1 is fully completed. Pashler (1992) used the analogy of a bank teller to describe the *Central Bottleneck Theory*. When a bank teller is faced with two customers almost at the same time, the customer that arrives later will experience the so-called *bottleneck delay*.

PRP effect was found to be existent in different paradigms incorporating discrete choice tasks (e.g., *left* and *right* hand response in a two-choice flanker task). For example, in an experiment conducted by Luck (1998), in which the participants were required to make two button-press responses for two consecutive stimuli (i.e., one to indicate the color of the first stimulus, and one to indicate the form of the second stimulus), a longer response time for the second stimulus was observed during the shorter SOA. In another experiment conducted by Tombu and Jolliceur (2002), participants were assigned to two tasks, the first task being a tone discrimination task and the second task being a shape-matching task. Different SOA was introduced and a response time delay caused by PRP effect was observed for the second task. There are many other studies investigating PRP effect, mainly debating whether this effect was caused by central capacity sharing, central bottleneck, or a combination of both.

Until recently, researchers utilizing PRP paradigm as a mean to investigate the human cognitive architecture have been incorporating two discrete choice tasks, particularly focusing on response time for the second task. Thus, here the overlap of decision processes was investigated. However, there have been no studies that investigated PRP effect in a subsequent *continuous tasks* like force production, which requires continuous motor control in its execution. The existence of a PRP effect in such subsequent tasks might modulate the information processing stages involved in the first and second tasks differently, in comparison to those involved in two discrete choice tasks. In the case of two successive simple force production tasks, more complex cognitive motor controlling processes are likely to be involved (i.e., to simultaneously execute a force production task and continuously monitor the task's performance), as opposed to the discrete choice task. For instance, in discrete choice tasks, the response selection process is often regarded as the main stage of the whole task execution, that is mainly processed during the *Central Processing* stage. However, in a continuous task like force production, the 'main task' includes not only a decision of the amount of force needed to be produced, but also other processes like: deciding how long the task execution is going to last (i.e., decision regarding how much time is needed to reach the Peak Force), as well as how many force units (see *PFUM*; Ulrich & Wings, 1991) are going to be deployed during the whole time period until the Peak Force is reached – which needs to be continuously monitored to ensure that the response force produced is as accurate as possible. That being said, the increased complexity of a continuous task might inflict further consequences – beside the prolonged response time for the second task, and these consequences might be observable not only on a behavioral level, but also on neural level.

1.4 Recent Developments on Force Production Monitoring

Despite the fact that our everyday activities usually consist of continuous movements, most of the studies in the area of action monitoring investigated tasks with *discrete* response parameters. This means, the response alternatives were clearly separated by effectors (e.g., *left* and *right* hand responses in two-choice flanker tasks, for *review* see Gehring et al., 2012). However, such binary left-or-right responses (e.g., Armbrecht et al., 2013; Stahl, 2010) are likely to be less complicated (in term of its motor execution) than a continuous force response which requires precise monitoring in a specific time interval, and are consequently easier to monitor.

The first to investigate action monitoring in a performance task with a continuous response parameter (i.e., the response force) in a task where the participants were required to produce a certain amount of force by pressing a button was De Bruijn et al. (2003). A mathematical model which accounted the relevant aspects of force production itself is called the *Parallel Force Unit Model* (Ulrich & Wing, 1991). In their study, the authors explained that two mechanisms determine the amount of force produced by the muscles: (1) the number of recruited force-producing motor units (so called *force units*; i.e., the more force units are activated, the higher the resultant response force; note, one motor unit consists of the muscle fibres activated by one motor neuron); and (2) the duration of force unit activation (i.e., the longer the force units are activated, the higher the resultant response force).

In De Bruijn's study on force production monitoring, the participants pressed a key such that the maximum of the produced force (peak force, PF) reached an amount within an individually defined range in one of two force conditions: low target force (e.g., 250–549 cN) and high target force (e.g., 550–850 cN). Thus, all responses with a force above and below these ranges are incorrect, which means there is not just one incorrect outcome but an unlimited number (for the given example, PF < 250 cN or PF > 850cN), which might explain why it is more difficult to monitor the quality of a continuous response parameter compared to binary left-or-right responses (e.g., Armbrecht et al., 2013; Stahl, 2010). Interestingly, De Bruijn et al. (2003) showed that for the *correct* responses, higher CRN amplitudes were observed during the *high target force* condition, compared to the *low target force* condition. This indicates that the action monitoring system was also sensitive to the magnitude of force required. Furthermore, the Ne/ERN amplitude was increased only for force selection errors, but not for force execution errors. A *force selection error* means that, for example, in a low force condition a high force was selected (e.g., higher than 549 cN in the above-mentioned *low target force* condition), whereas a *force execution error* occurs when participants aimed for the correct *low force* range but produced an even lower force. Based on these findings, Armbrecht et al. (2013) reasoned that the force selection errors might not reflect the actual errors *per se*, but could be directly linked to controlling different aspects of the specific response dynamics (i.e., preparing the response force, or initiating the timing in response execution). To investigate this possibility, they assessed the effects of two response parameters on the Ne/ERN and CRN independently: (a) *peak response force* (PF) and (b) *Time-To-Peak* (TTP, i.e. the time period between response onset and maximum point of the force pulse). Participants were instructed to produce four types of isometric force pulses (i.e., high PF with a short TTP, high PF with a long TTP, low PF with a short TTP, and low PF with a long TTP). Independent of response accuracy, the amplitudes for both Ne/ERN and CRN were higher in the high target force condition compared to the low target force condition (similar to de Bruijn et al.'s findings). Interestingly, Armbrecht et al. also found significantly reduced amplitudes for both Ne/ERN and CRN for long TTP compared to short TTP, independent of response accuracy. However, they found no error-related effects (which would correspond to *force execution errors* in Bruijn et al.'s terminology) for Ne/ERN and CRN. Taken together, these studies provided evidence that Ne/ERN and CRN are affected by continuous response parameters (peak force and TTP) as well as by error-specific variations, but the functional meaning of the Ne/ERN and CRN in these tasks was still not clear.

Notably, neither de Bruijn et al. (2003) nor Armbrecht et al. (2013) found clear neural evidence of error processing during *force execution*. However, a previous study on pianists, who are highly trained in force execution, showed that expertise is linked to more precise response dynamics of finger presses (Parlitz et al., 1998). The results suggest that the action monitoring system is indeed capable of tracking response force during execution. Taking these results into consideration, the absence of effects during *force execution* stage in de Bruijn et al. (2003) and Armbrecht et al. (2013) studies could be due to the lack of substantial training for the participants.

1.5 The Current Research

The studies outlined above found variations depending on the force levels, yet it remains unclear how these differences would unfold during the *force execution* stage. Understanding how force execution unfolds was the focus of the first study. By modifying the force production paradigms used in the previous studies, an attempt to gather more information about response dynamics in a simple one-dimensional force production, as well as to elucidate the time course of information relevant to the monitoring processes was made. To do this, not only classical ERP components but also a multivariate approach was used to find pattern in the brain activity to distinguish two different experimental conditions (i.e., different force conditions and response quality), and to predict continuous cognitive variables such as response force parameters (i.e., Peak Force and Time-to-Peak).

The second study was designed to investigate the course of two subsequent simple one-dimensional force tasks. In such subsequent tasks, depending on the length of the interval between the two tasks, the information processing stages (which consist of different processes such as response planning and error monitoring) for the second task might be somehow impaired or less efficient (i.e., response time for the second task is prolonged when the interval between the two tasks is very short). Such impairment on the second task's processing has been widely known as a 'PRP effect', and has been observed in studies using two subsequent discrete tasks such as flanker tasks. It is particularly interesting to see how varying the length of interval between subsequent force tasks would modulate the processing of the latter task. Thus, an adapted PRP paradigm for two subsequent force production tasks was used in the second study, to understand how one or more parts in the information processing stages of a continuous task such as force production is affected by different Response-Stimulus Interval / RSI (i.e., an adapted version of *Stimulus Onset Asynchrony* / *SOA* used in the second study's paradigm). A further challenge would be, to explain where in the information processing stages the PRP effect comes into play, and how it modulates the second task's response parameters on a behavioural as well as neural level.

II. Empirical Evidence

2.1 Study 1

2.1.1 Objective of the Study

The first study was focused on investigating action monitoring during *force execution* in a simple one-dimensional force production task. The general idea was to learn more about response dynamics in force production, and to elucidate when information relevant to the monitoring processes (i.e., response force parameters, error specific information) becomes available and was reflected in brain activity. For this reason, two approaches were used: (1) classical ERP components, and (2) a chronometric, multivariate approach (e.g., Bode & Stahl, 2014). Using the latter method provides an advantage in addition to the classical univariate approach: it allows for identifying the onset of the availability of specific information regarding response parameters (i.e. response quality, force magnitude, TTP) in distributed patterns of brain activity. It can also be used – by the means of multivariate regression – to predict single-trial response parameters from brain activity patterns. This multivariate approach has been used successfully in detecting movement intentions (Jochumsen et al., 2016), classifying different movements when force and speed are varied (Jochumsen et al., 2013), as well as error detection before an overt response in a digit-flanker task (Bode & Stahl, 2014).

In order to focus on *force execution,* a modified version of force production paradigms from preceding studies (de Bruijn et al., 2003; Armbrecht et al., 2013) was utilized by using just *one* target force range (with two possible *force execution errors*: responses above the *target force* range, here referred to as the *"too high condition"* and below the *target force* range, referred to as the *"too low condition*"). The decision to incorporate only one target force range was made to fully eliminate *force selection errors* and to further diminish the need of substantial training of different target forces. To minimize confounding factors which might be introduced by further TTP variations, participants were further trained to produce only one specific brief force pulse with a short TTP in each trial. Thus, participants could not confuse the required target range with the target pulse length. In each trial, force parameters could be defined in terms of PF and TTP, while errors could only be defined as missing the target force range (i.e. under-producing or over-producing the required force).

The modified paradigm used in this study allows for linking the classical ERP components to different aspects of response dynamics. First, several contrasting hypotheses for Ne/ERN as well as CRN were tested: if the Ne/ERN merely reflects error monitoring (but not force magnitude monitoring), no difference between the Ne/ERN in the two error conditions (*too low* and *too high*) would be expected. However, for both error conditions, the Ne/ERNs should be larger than the CRN. If the observed components are sensitive to the force magnitude but not to the *correctness* of the response, as suggested by Armbrecht et al. (2013), the components' amplitudes should be scaled accordingly to the conditions (i.e., smallest when under-producing the force, medium when reaching the target force range, and highest when over-producing the force). If the components happened to be reflect both error processing and force production monitoring, identifying clear differences between conditions might be more difficult.

Secondly, the Pe component was investigated. Although Pe has been neglected in recent force monitoring studies, this component establishes an interesting indicator of aware error detection (e.g., Nieuwenhuis et al., 2001) as well as the accumulation of error evidence (e.g., Steinhauser & Yeung, 2012). The Pe usually peaks around 300 ms after response, which in this case after the peak force is usually reached, therefore making the Pe an interesting measure in term of force-related error detection.

The FRN was the third ERP component investigated in this study, since this component reflects externally-induced mechanism of error processing. According to the *first-indicator hypothesis* (see Holroyd & Coles, 2002), if an error is already detected during the time of responding, a feedback would not provide additional information, which means that error-related effects should not be reflected in the FRN (but only in the Ne/ERN). On the other hand, if errors were not detected at the time of responding, the (external) feedback would serve as the first indicator of error commission, which means that the FRN (but not the Ne/ERN) should reflect error-related information (Stahl, 2010).

The second motivation for this first study was to better understand how *force execution* process unfolds. To do this, a chronometric multivariate approach (MVPA; see Bode et al., 2012; Bode & Stahl, 2014) was used. This multivariate approach takes into account not only spatial but also temporal aspects of force production. The goal was to investigate if information related to a force response (i.e. response quality / *correctness*, types of error, force magnitude and TTP) could be foreshadowed from the spatially-distributed ERP signals in small, consecutive time-windows while the response was planned and subsequently executed. This technique allows for detecting *when* (before and/or after the response onset) information that predicts the response dynamics of single trial response force parameters (i.e., PF, TTP) becomes available in the brain activity pattern, as well as *whether* the target force range will be missed (errors vs. correct responses), or, in case where an error is committed, force will be under- or overproduced. These analyses might, beyond clarifying the role of ERP components in force monitoring, further explain how a simple 1-dimensional response force is produced, i.e. how early specific response information is set up and locked in, which would be essential for early error processing in the current task.

2.1.2 Method

2.1.2.1 Participants

Seventy-eight participants (46 females) from the University of Cologne participated in this study. Due to insufficient number of errors (less than 10 trials per condition) and technical artefacts, data from 30 participants were excluded. The analyses reported here were based on data from the remaining 48 participants (33 females, age range: 17 to 44 years; mean ± SD: 25.47 ± 5.44 years). All participants had normal or corrected-to-normal vision, and were right handed. Informed consent was obtained from all participants, and they were rewarded with course credit for their participation.

2.1.2.2 Apparatus

A custom-made force-sensitive key, mounted on a board which provided full forearm support (see Figure 4), was used to record the behavioural data. Participants initiated a response by briefly pressing the force key with their right index finger. The response force was measured by a strain gauge attached on the fixed end of a leaf spring, which was held by an adjustable clamp at one end, leaving the other end free for the participant to press. When participants pressed the free end of the force key, an analogue electrical signal which corresponds to the exerted force was produced. This signal was recorded with a sampling rate of 500 Hz. A chin rest with an adjustable height was used to maintain a constant and stable posture while keeping a distance of 56 cm from the monitor.

2.1.2.3 Maximum Voluntary Force

Before the experiment, participant's *maximum voluntary force* (MVF) was assessed. This assessment was necessary to determine the individual force ranges (defined in % MVF; see below) as MVF varies across participants. The participants received an instruction to press the force key with their right index fingers as hard as possible without moving the forearm. This procedure was repeated ten times in a row. A start signal initiated each of the ten key presses. The individual MVF was calculated by averaging the PF of the last seven key presses (e.g., Armbrecht et al., 2013). Five force ranges relative to the individuals' MVF were then defined (*target range*: 46–54% MVF; *too high*: > 60% MVF; *slightly too high*: 54–60% MVF, *slightly too low*: 40–46% MVF; *too low* < 40% MVF).

2.1.2.4 Experimental Task

Participants were tested individually in a 50-minute experimental session. The experiment comprised six blocks, each consisting of 44 trials. Each trial began with a presentation of a white square in the middle of the screen, serving as a start signal. Participants were then required to produce a brisk, isometric force pulse with their right index finger. To minimize TTP variability, a timed-feedback presentation was used (the white square became red if the participants failed to reach the peak force after 180 ms) in the first block. In all blocks, force feedback was presented for 700 ms after a response was made, to indicate whether the participants successfully reached the target range (see Figure 3): The green area - located in the middle of the *force ruler* corresponded to the 'correct' area. The upper red area corresponded to the *too high* condition, and the bottom red area corresponded to the *too low* condition. The yellow areas represented the '*slightly too high'* and '*slightly too low'* conditions (the yellow areas were not included as separate conditions in the final analyses). A white cursor (see Figure 3B) represented the force level that was produced in the respective trial. The feedback presentation was followed by a black screen with a randomly jittered inter-trial duration (300-600 ms).

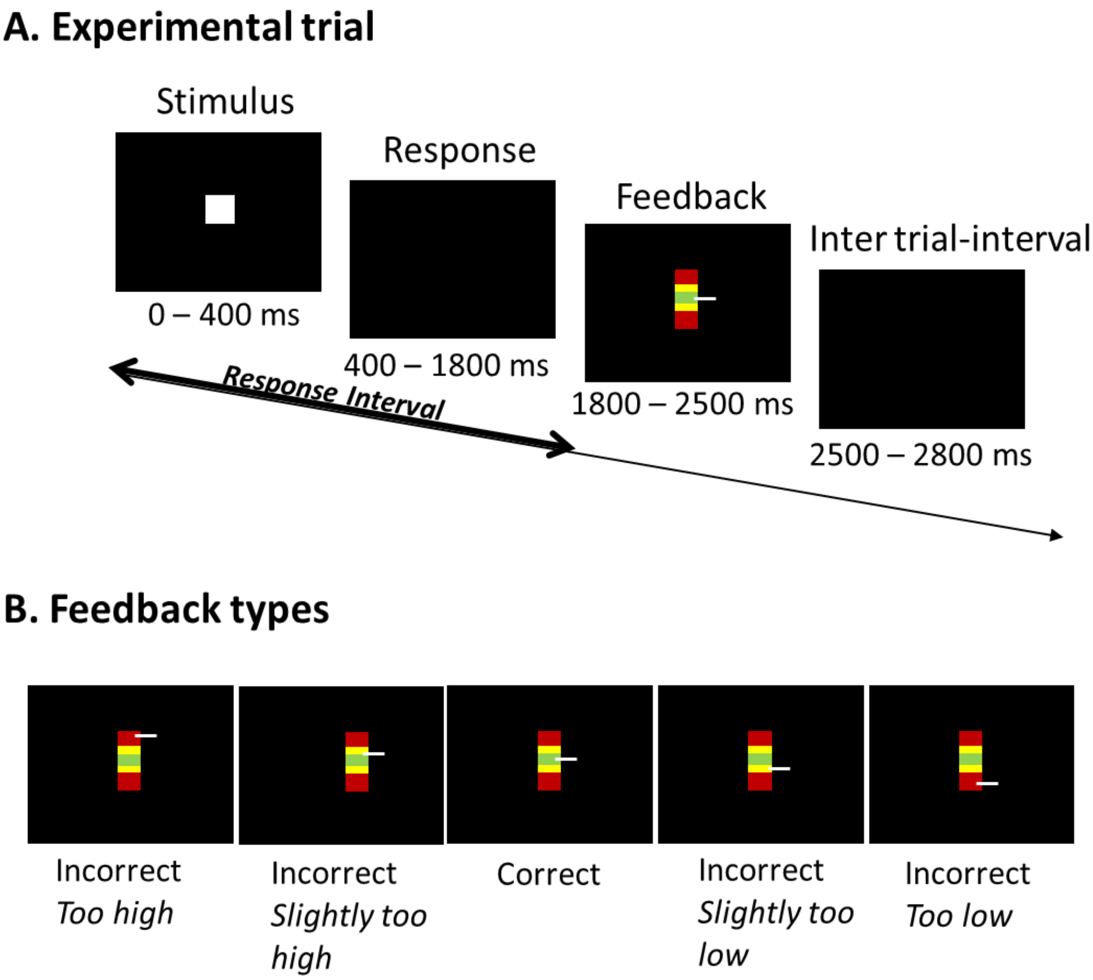


Figure 3. (A) Time course of each trial and (B) all feedback types in the force task. A response was indicated by a key press; the feedback was presented according to the force produced in the respective trial.

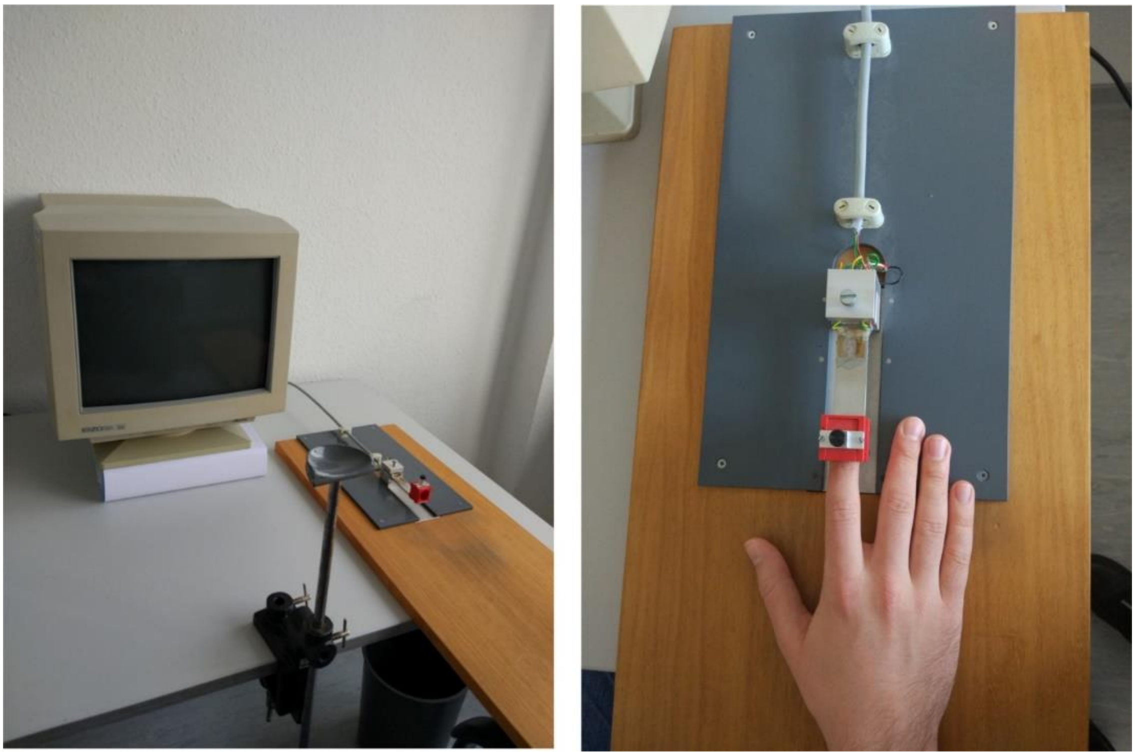


Figure 4. Experimental Setup: a strain gauge attached on the fixed end of a leaf spring (measurements: 110 x 19 x 2 mm) held by an adjustable clamp at one end of the key was used to measure response force; participant pressed the free end of the key to produce an analogue electrical signal corresponding to the exerted force.

2.1.2.5 Data Acquisition

Behavioural Data

Response time (RT) was defined as the first time point when the participant's response force exceeded 50 cN (i.e. response onset), measured from the stimulus onset. *Peak force* (*PF*) was defined as the maximum of the force (in cN) of a single trial force pulse. *Time-to-Peak force* (TTP) was defined as the time point (measured from response onset) at which the peak force was reached. Frequency of correct responses, frequency of *too high* responses, frequency of *too low* responses, mean RT, mean PF, and mean TTP were determined separately for each condition and each participant.

Electrophysiological Data

EEG data were obtained from 61 scalp electrode sites according to the standard international 10–20 system (Jasper, 1958). The active Ag/AgCl electrodes (actiCAP, Brain Products) were referenced against the left mastoid. The vertical electrooculogram (EOG) was recorded from an electrode infraorbital to the left eye and the horizontal EOG was recorded at the outer canthi of each eye. The EEG was continuously recorded at a sampling rate of 500 Hz using BrainAmp DC (Brain Products).

The electrophysiological data were time-locked to (a) the response onset; (b) the *Time-To-Peak*; and (c) the feedback onset. In the case of *response-locked* and *feedback-locked* analyses, data were epoched ranging from 100 ms before until 600 ms after response (or feedback) onset. For the *Time-To-Peak-locked* analyses, data were epoched ranging from 300 ms before until 300 ms after peak force was reached. Baseline correction was performed with the period of 100 ms before response/feedback onset for the *response-locked* and *feedback-locked analyses,* and with the period of 100 ms before stimulus onset for *TTP-locked analyses*. An ocular correction algorithm was applied in order to reduce the impact of eye movements (Gratton, Coles, & Donchin, 1993), followed by a second baseline correction. Afterwards, an artefact-rejection procedure was carried out to eliminate contaminated trials exceeding maximum/minimum amplitudes of ± 100 μV. The remaining trials were averaged. A current source density (CSD) analysis was then performed on the averaged ERP waveforms. This analysis accounted for the curvature of the head using a spline algorithm (Perrin et al., 1989), and was performed to reduce the effect of neighbouring currents. The CSD signals (order of splines = 4; lambda = 10^{-5}; maximal degree of Legendre polynomials = 10) were computed for each electrode site by taking the second derivative of the distribution of the voltage over the scalp. Lastly, the ERP components (Ne/ERN, CRN, Pe/Pc for correct trials and FRN; both peak amplitudes and mean amplitudes as the standard indicator for the area under the curve in the defined time ranges – hereafter referred to as ‘area’) were determined separately from the individual mean CSD-ERP waveforms at the electrode sites FCz and Cz. Ne/ERN (CRN) was defined as the most negative peak in a time window ranging from 0–180 ms after response onset at electrode site FCz. Pe (Pc) was defined as the most positive peak in a time window ranging from 150–300 ms after response onset at electrode site Cz. Finally, the FRN was defined as the most negative peak in a time window ranging from 150–250 ms after feedback onset at electrode sites Cz and FCz. The selected time windows and electrode sites for each component were in line with previous studies (e.g., de Bruijn et al., 2003; Falkenstein et al., 2000).

2.1.2.6 Univariate Statistical Analyses

Several repeated-measures analysis of variance (ANOVA) were conducted for the within-subject factor *force range* (*correct, too high, too low*) for all behavioural measures (RT, PF, TTP, $\Delta RT_{post,}$ ΔPF_{post}) using SPSS 23. Due to low number of error trials (less than 6 trials), the *slightly too high* condition and *slightly too low* condition were not incorporated for the separate analyses (but the trials were included for the regression approach of MVPA). Separate ANOVAs with the within-subject factor *force range* were further performed for the respective ERP components' peak amplitudes (Ne/ERN, CRN, Pc, Pe) and the components' area measures. A two-way ANOVA with the within-subject factors *force range* (*correct, too high, too low*) and *electrode site* (FCz and Cz) was performed for the FRN peak amplitude. The electrode factor was used here as the source of the component was less clear compared to the other components. Significant ANOVA results were followed up using Bonferroni adjusted post-hoc tests. Level of significance were adjusted using Geisser and Greenhouse (1958) correction in case the sphericity assumption was violated. Effect sizes are reported in terms of partial eta² (η_p^2).

2.1.2.7 Multivariate Pattern Classification Analysis

Multivariate pattern classification analysis (MVPA) was used to find the earliest time point after stimulus presentation that allowed for decoding the response outcome from distributed spatio-temporal patterns of ERPs. In this study, two types of MVPA analyses were conducted using the Decision Decoding ToolBOX (Bode et al., 2019). For the first set of analyses, *support vector machine classification* (SVC) was used in a moving analysis window approach to decode the following conditions in the force task: (1) *correct* vs. *too high* condition; (2) *correct* vs. *too low* condition; (3) *too high* vs. *too low* condition. Additionally, *too high* and *too low* conditions were aggregated to one condition which was referred to as the *incorrect* condition, and an additional SVC analysis for *incorrect* vs. *correct* was conducted. As the number of trials in the *too high* and *too low* condition was imbalance (with the ratio of 2:3), the number of trials for these two conditions was balanced for both the training and test data sets of the *incorrect* condition. In all cases, the moving analysis windows were applied to the brain activity pattern before and after the response onset. For the second set of analyses – aimed at identifying the first time point at which the response parameters (PF, TTP) were decodable –*support vector regression* (SVR) was used, which allows the decoding of continuous outcome variables.

Classification Analyses

For the first set of classification analyses using SVC, each participant's artefact-cleaned data was sorted into three groups with respect to the force produced: (1) *correct*; (2) *too high*; (3) *too low*. For an additional *correct* vs. *incorrect* analysis, group 2 (*too high*) and group 3 (*too low*) were combined into one group (*incorrect* condition). The EEG epochs (for all groups) were time-locked to the response onset and (additionally) the time point when the peak force was reached (TTP). Then, six classification analyses were conducted to analyses three pairwise classification analyses for both response-locked and TTP-locked data, respectively. For the main classification analyses (response-locked), data starting from 150 ms before response onset was included since Bode and Stahl (2014) were able to predict whether responses would be erroneous or correct from about 100 ms before response onset. As the task was to reach a specific maximum force, and the TTP point marked the end point of the force production phase, it was of interest to decode information in neural signals leading up to this time point, mirroring the classical ERP analyses. For these classification analyses, data starting from 300 ms before the peak force and after the peak force was reached were included.

For each individual pattern classification analysis (*correct* vs. *too high*; *correct* vs. *too low*; *too low* vs. *too high; correct* vs. *incorrect*) the following steps were performed. A non-overlapping *spatio-temporal* analysis time-window of 10 ms that contained 5 data points for each of all 61 channels was used, covering the entire epoch (cf., Bode & Stahl, 2014). For each trial, all data points included in this window were transformed into vectors, to represent the *spatio-temporal* patterns associated with each condition. These were then randomly assigned to ten separate sets. The linear classifier (using the default regularisation parameter C = 1) was trained on vectors from the two conditions of interest by interfacing with the LIBSVM toolbox (Chang & Lin, 2011). Importantly, only 90% of the data (i.e., 9 of the 10 sets) were randomly drawn and were used to train the classifier. Based on these exemplars, the classifier estimated a decision boundary that optimally separated exemplars from the two classes (i.e. categories). The vectors from the remaining independent 10% of the data were subsequently used for testing the classifier (Bode et al., 2019). Note that in order to avoid biased results because of an imbalanced trial numbers (i.e., samples), the trial numbers for conditions before dividing the data into training and test sets were balanced. Note that only the smaller number of trials of the two conditions were used, and the trials of the condition with larger trial numbers were randomly drawn to match this number. This means, there was an equal number of trials in both the training set and the test set, and the number of exemplars was always the same between conditions.

The percentage of correct classifications (decoding accuracy) is based on the estimated decision boundary served as the outcome measure. Above chance classification indicates that the data from the respective analysis time window contained information about the two conditions, while chance-level classification (50% with two classes) suggests no evidence for such information. In order to minimise the risk of false positive results when determining the decoding accuracy, the classification process was first repeated using a 10-fold cross-validation procedure in which each set containing 10% of the data served as *test data* once while the classifier was re-trained on the remaining 90% of the data, until all of the sets were independently used for testing once. In addition, to avoid potential drawing biases, all ten cross-validation steps were then repeated ten times in an identical fashion, but with newly randomly drawn 10 sets for training and testing, resulting in a total of 100 analyses. The average of all of these analysis steps constituted the estimate of the individual classification accuracy for the respective analysis time window.

Finally, statistical testing was performed at group-level. For this, a series of Bonferroni-corrected t-tests for each (10 ms) analysis time window were conducted to test the performance of the classifier against empirical chance results from the near-identical shuffled-labels analysis (Bode et al., 2019), in which the same data and the same labels were used in the same number of cross-classification steps (and iterations thereof), but the assignment of labels to data was randomly assigned for each step. This means, both the real and the empirical chance distributions were composed of the average accuracies (or Fisher-Z transformed correlation coefficients for SVR) across 100 analyses per participant (10 x 10 cross-validation steps). The decision to use the empirical chance distributions was made because it provided a stricter test than testing against the theoretical chance level of 50% (Bode et al., 2012; Bode & Stahl, 2014).

Support Vector Regression (SVR)

For this analysis, all artefact-cleaned data (without further grouping into any condition) were included. A slightly longer interval of the same EEG epochs (response-locked, -400 ms to 300 ms related to response onset) and the same set of parameters for the *spatio-temporal* analysis (non-overlapping analysis time windows of 10 ms that included 5 data points from each of the 61 channels) were used as for the classification analysis. The decision to include data points starting at 400 ms before response onset for SVR was made because if evidence regarding erroneous response was present in the brain already around 100 ms before response onset (see Bode & Stahl, 2014), information regarding response related parameters (i.e. PF and TTP) might be present in the brain activity pattern even earlier. The model was again trained on vectors from all data using LIBSVM (we used the default parameters epsilon-SVR = 3 and C = 1). To estimate the regression model, data from all trials (including the *slightly high force/slightly low force conditions*) were again, randomly divided into 10 equally-sized sets. Out of these, 9 (90% of all data) sets were randomly drawn and used to train the model, while the left-out one (10% of all data) was used for testing. As before, a 10-fold-cross validation procedure was then applied for which each data set was used for testing once and the other 9 were used for training. The same strict 10-times repetition of the entire cross-validated analysis was conducted as described above. The only difference to the classification analysis was that the results of SVR are individual correlation coefficients between the *predicted* variable (PF or TTP) based on the regression model and their *true* values, averaged across all iterations for each given time window. a Fisher-Z transformation for the correlation coefficients was conducted, and the final measure was one average coefficient per participant per analysis time window, reflecting information regarding the condition of interest (e.g., Bode et al., 2014). In other words, a significant coefficient means that information regarding PF (or TTP) was represented in brain activity in the respective analysis time window. The same analysis was then conducted with randomly shuffled labels (i.e., PF or TTP values) for each participant and each analysis time window to obtain an empirical distribution of regression results under the null hypothesis for each analysis step. Group level Bonferroni-corrected t-tests were then used to compare the real empirical results with the shuffled label SVR results for each time window.

2.1.3 Results

2.1.3.1 Behavioural Data

The first set of analyses was done to investigate whether responses for the different peak force ranges classified as *too high*, *correct* and *too low* also differed with respect to TTP. The respective ANOVA showed that Force Range had a significant effect, $F(2,47) = 22.79$, $p < 0.001$, $\eta_p^2 = 0.245$. The longest TTP was observed in the *too high* condition (252.74 ± 13.42 ms), followed by the *correct* condition (245.38 ± 15.56 ms) and the *too low* condition (212.95 ± 12.53 ms). Follow-up post-hoc tests confirmed significant differences between the *correct* and the *too low* condition ($p < 0.001$), as well as between the *too high* condition and the *too low* condition ($p < 0.001$), but not for the *too high* condition and the *correct* condition ($p = 0.999$). No significant Force Range effect was observed for RT, $F(2,47) = 0.325$, $p = 0.723$. Distribution of response types was also investigated (see Figure 5), and the correct responses were found to be the most frequent (46.44 ± 1.39%), followed by the *too low* condition (18.82 ± 1.30%) and the *too high* condition (12.81 ± 0.82%).

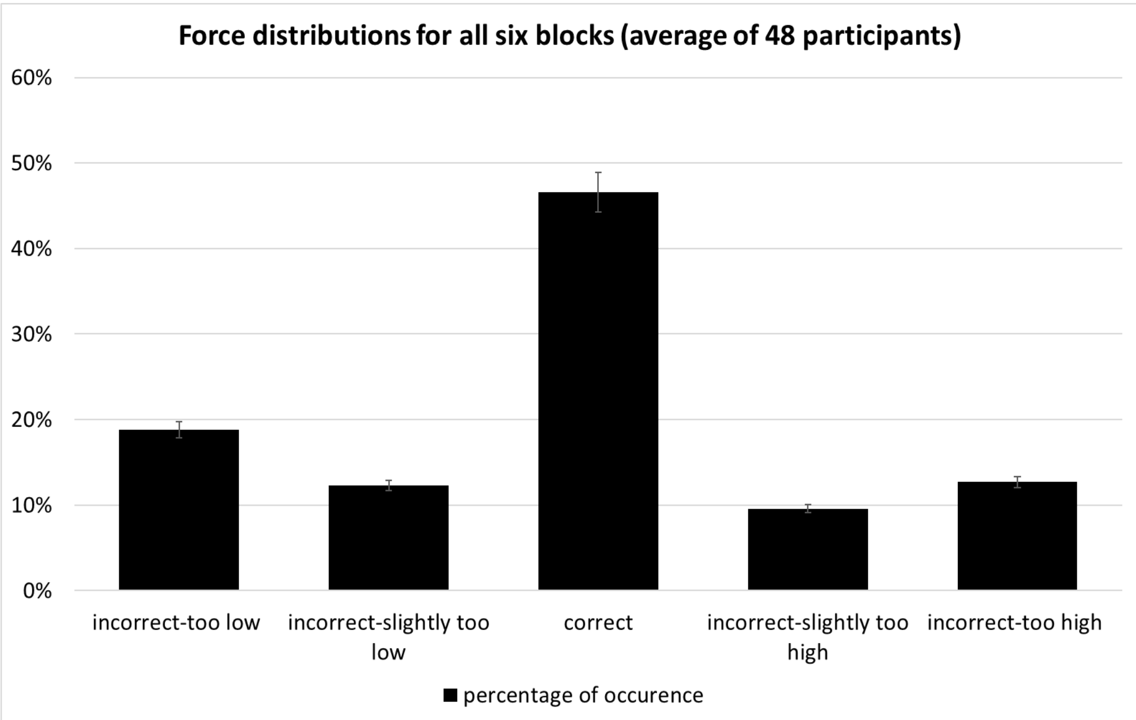


Figure 5. Average force distributions for all six blocks (N=48)

2.1.3.2 Electrophysiological Data

Response-locked and TTP-locked Averages

A series of ANOVAs was conducted, each testing the effect of the within-subjects factor Force Range (conditions: *too low, correct, too high*) on the peak amplitude in the time window of interest (specific for each ERP component) at the respective channels, as described in the method section. Identical analyses for the same components were conducted with data time-locked to the response and to the TTP, respectively. The Force Range did neither affect the response-locked Ne/ERN peak amplitude (CRN for the *correct* condition), $F(2,47) = 3.08$, $p = 0.050$, $\eta_p^2 = 0.062$, nor Ne/ERN area, $F(2,47) = 0.86$, $p = 0.424$ (see Figure 6A). The additional TTP-locked analyses for Ne/ERN (see Figure 6B) showed a similar pattern of results. Neither Ne/ERN peak amplitude, $F(2,47) = 0.59$, $p = 0.554$, $\eta_p^2 = 0.025$, nor Ne/ERN area, $F(2,47) = 1.68$, $p = 0.197$, $\eta_p^2 = 0.068$, showed significant effects.

Significant effects of Force Range were detected for both response-locked Pe peak amplitude, $F(2,47) = 3.77$, $p = 0.030$, $\eta_p^2 = 0.141$, and Pe area, $F(2,47) = 6.98$, $p = 0.002$, $\eta_p^2 = 0.233$, at Cz (Figure 6C). Post-hoc tests showed significant differences for Pe area between the *correct* condition (11.29 ± 1.40 $\mu V/cm^2$) and the *too high* condition (15.42 ± 2.26 $\mu V/cm^2$, $p = 0.017$), as well as between the *too low* condition (9.95 ± 1.26 $\mu V/cm^2$) and the *too high* condition (15.42 ± 2.26 $\mu V/cm^2$, $p \leq 0.001$). A near-identical pattern of results for the TTP-locked analyses for Pe peak amplitude and areas was also observed (see Figure 6D). Significant effects of Force Range were observed for both Pe peak amplitude, $F(2,47) = 7.521$, $p < 0.001$, $\eta_p^2 = 0.246$, and Pe area, $F(2,47) = 15.569$, $p < 0.001$, $\eta_p^2 = 0.471$, at Cz. Post-hoc tests showed significant differences for Pe area between the *correct* condition (10.36 ± 1.27 $\mu V/cm^2$) and the *too high* condition (16.72 ± 1.85 $\mu V/cm^2$, $p < 0.001$), as well as between the *too low* condition (9.87 ± 1.24 $\mu V/cm^2$) and the *too high* condition (16.72 ± 1.85 $\mu V/cm^2$, $p \leq 0.001$). Highly similar post-hoc results pattern for Pe peak amplitude were observed for both response-locked and TTP-locked results. Additionally, response and time-to-peak locked analyses for *correct* vs. *incorrect* (combination of *too low* and *too high* trials) were also presented (see Figure 7) to give a better illustration regarding the general differences between the *correct and erroneous* conditions.

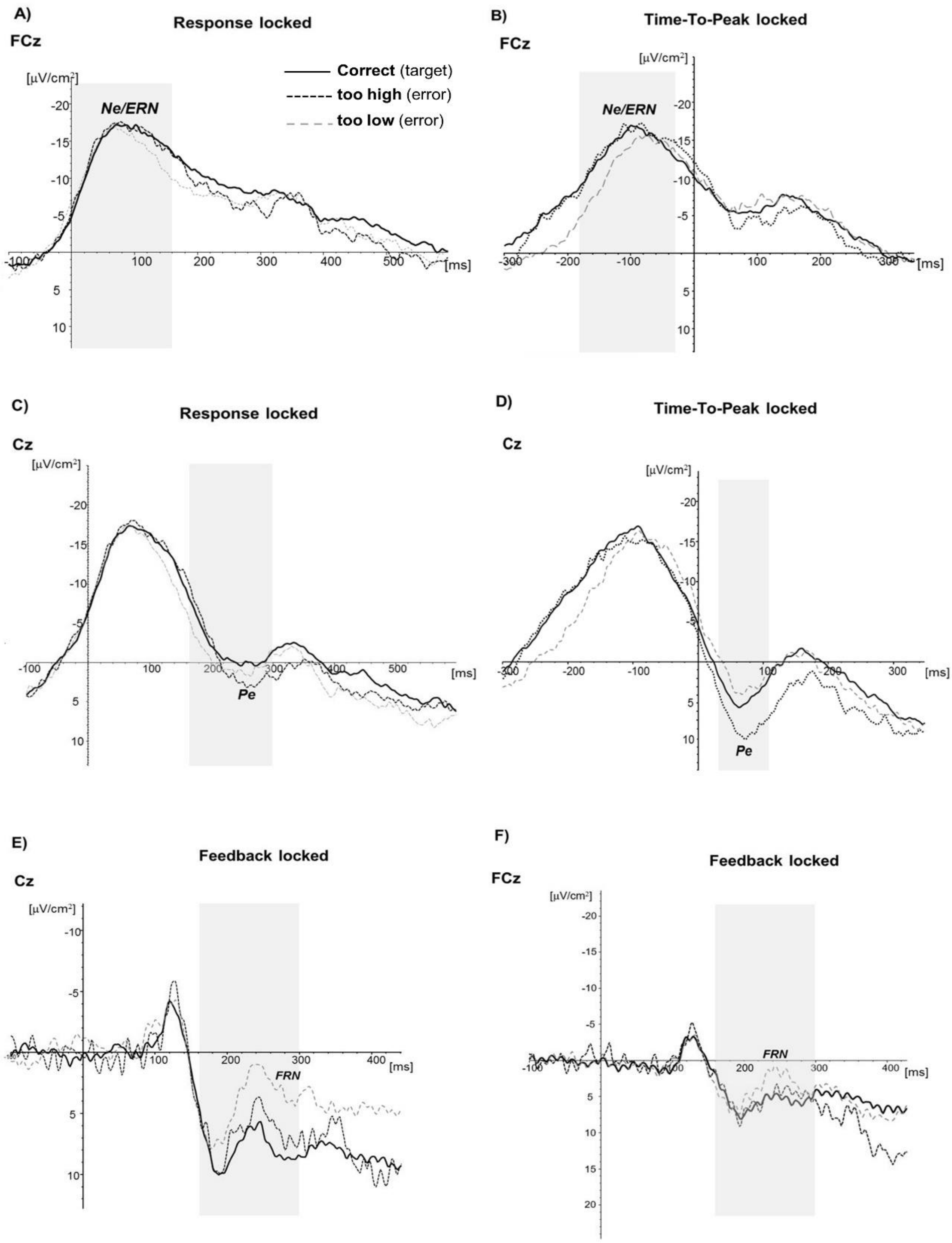


Figure 6. Averaged event-related potentials of the force task time-locked to the response onset (A, C), time-locked to the Time-To-Peak (B, D), and time-locked to the feedback-onset (E, F) for the two electrode sites of interest FCz (A, B, F) and Cz (C, D, E) for error–related negativity (Ne/ERN), error positivity (Pe) and feedback-related negativity (FRN).

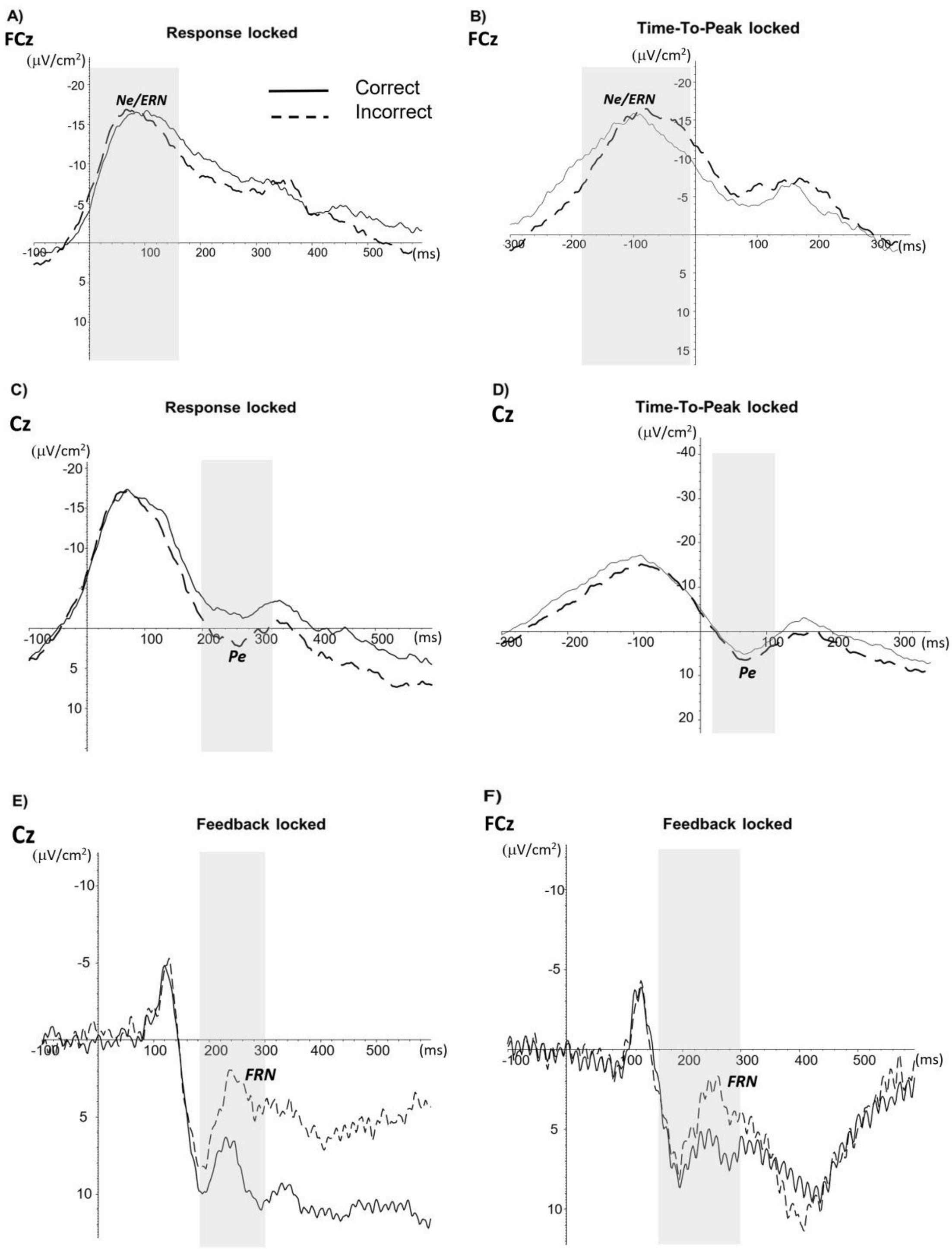


Figure 7. Averaged event-related potentials of the force task for **correct vs. incorrect** condition time-locked to the response onset (A, C), time-locked to the Time-To-Peak (B, D), and time-locked to the feedback-onset (E, F) for the two electrode sites of interest FCz (A, B, F) and Cz (C, D, E) for error–related negativity (Ne/ERN), error positivity (Pe) and feedback-related negativity (FRN).

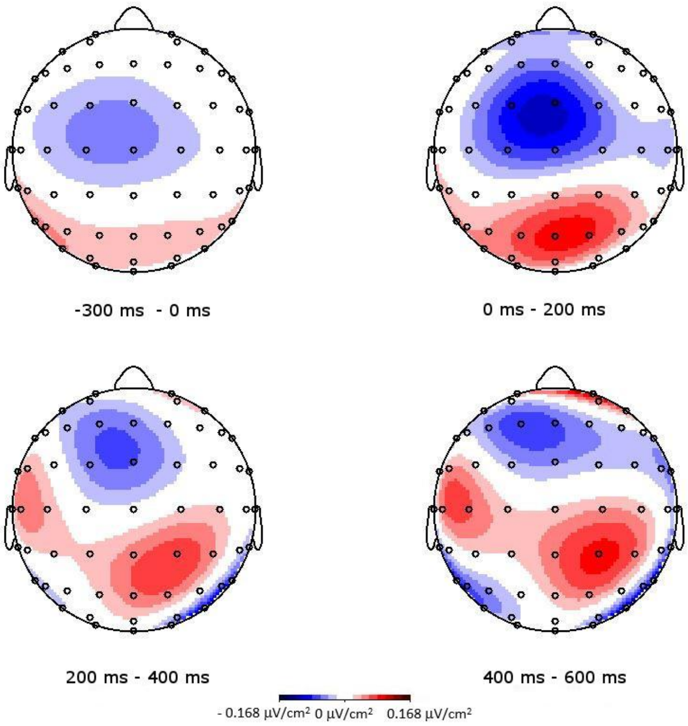


Figure 8. Topography distribution for the ERP signals

Feedback-locked Averages

The ANOVAs for FRN peak amplitude (see Figure 6E, F) revealed a significant effect of Force Range at Cz, $F(2,47) = 6.43$, $p = 0.002$, $\eta_p^2 = 0.120$. Post-hoc comparisons showed that FRN peak amplitude of the *too low* condition (-5.08 ± 1.95 µV/cm^2) was significantly more negative than of the *correct* condition (0.45 ± 1.88 µV/cm^2; $p = 0.002$). No significant differences were observed between the *too high* condition (-3.05 ± 1.87 µV/cm^2) and the other two conditions (*ps* > 0.10). A near-identical pattern of results was observed when the FCz electrode was used to determine the FRN. Here, FRN area was not investigated because for a relative negative component, the area is usually not very informative.

2.1.3.3 Multivariate Pattern Classification

To investigate *whether* error-specific information is predictable from the distributed ERP signal, and to identify *when* (i.e., time point(s)) this error specific information is decodable from brain activity, the first six sets of multivariate classification analyses were conducted. The first three classifications were conducted using response-locked ERP data (see Figure 9), and the other three were conducted using TTP-locked ERP data (see Figure 10). For each analysis, each pair of the defined force ranges were used as distinct classes for the classifier, each time using a moving-window approach with 10 ms analysis time windows: *correct* vs. *too high* (Figure 9A,10A), *correct* vs. *too low* (Figure 9B, 10B), and *too high* vs. *too low* (Figure 9C, 10C). We also conducted a *correct* vs. *incorrect* classification analysis (Figure 11) to investigate whether a more 'general' error-related processes could be reflected in the ERP signal.

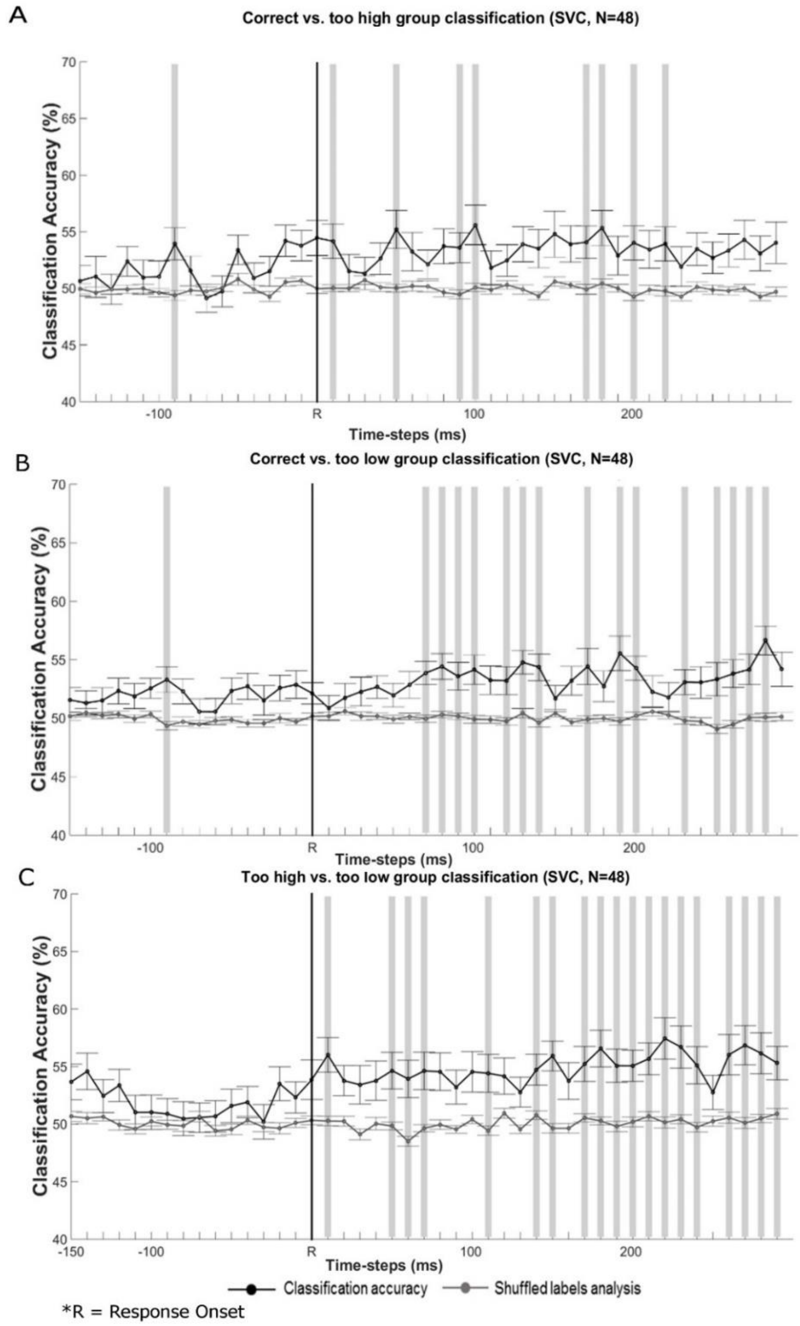


Figure 9. Classification accuracies of group classifications and permutation test results using multivariate pattern analysis (Support Vector Classification, SVC, A-C), time-locked to the Response Onset (R), for (A) correct vs. too high classification, (B) correct vs.

too low classification, (C) too high vs. too low classification; corrected for multiple comparisons (N = 48; p < .05; error bars indicate standard errors of the mean; grey time windows indicate significant accuracy).

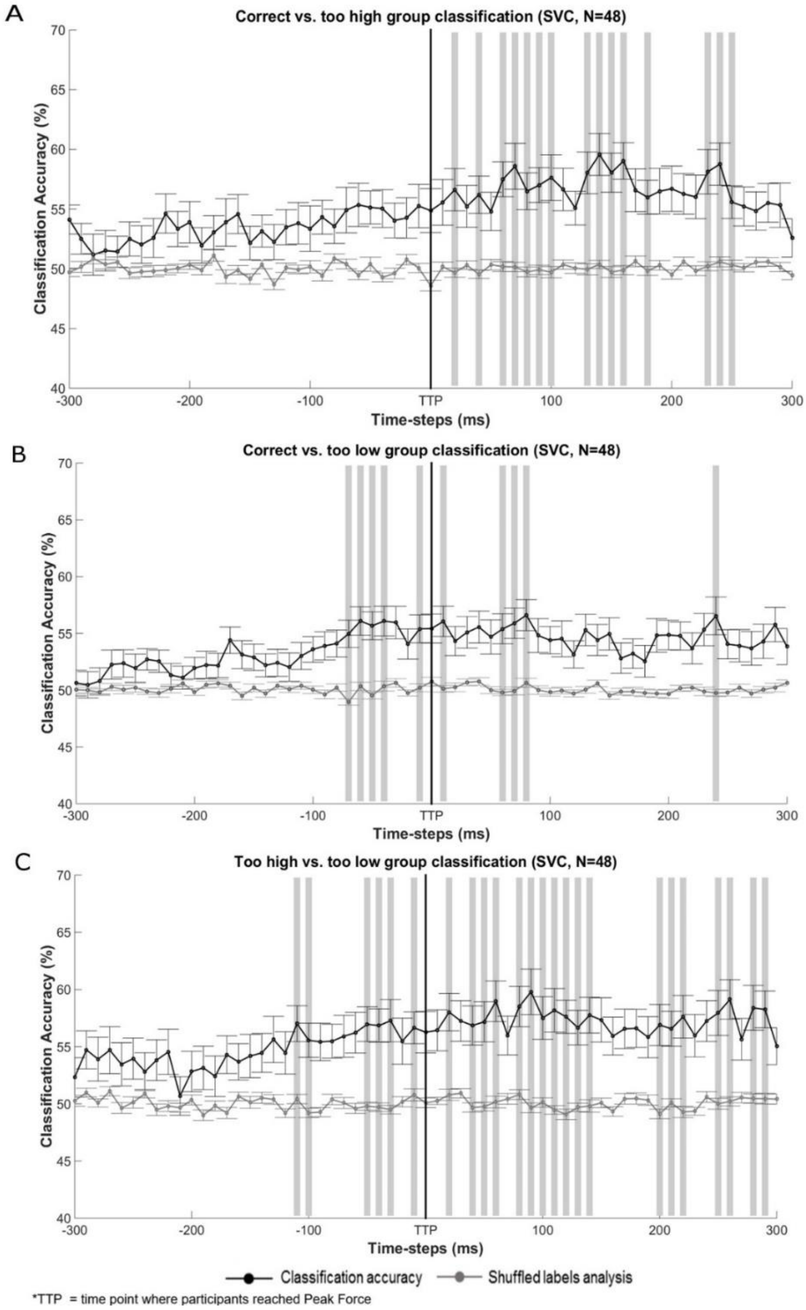
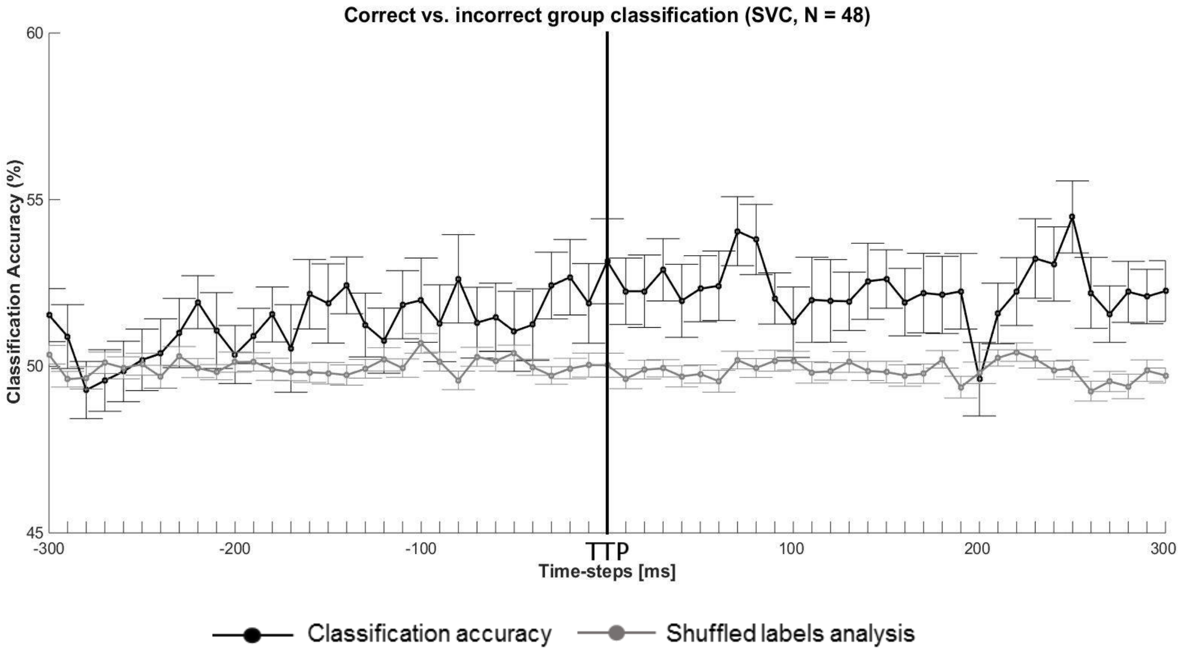


Figure 10. Classification accuracies of group classifications and permutation test results using multivariate pattern analysis (Support Vector Classification, SVC, A-C), time-locked to the Time-to-peak (TTP), for (A) correct vs. too high classification, (B) correct vs. too low classification, (C) too high vs. too low classification; corrected for multiple comparisons (N = 48; p < .05; error bars indicate standard errors of the mean; grey time windows indicate significant accuracy).

The main (response-locked) classification results (Figure 9) showed significant classification accuracies for all three analyses. For all three analyses, no evidence for substantial prediction before response onset was found (0 ms), but several significant time windows between response onset and the average TTP (ranging between 213 and 253 ms, see behavioural data) were identified, indicating that information about the quality of the response outcome was available in brain activity before the actual maximum force was reached. The additional TTP-locked classification analyses (Figure 10) confirmed that information regarding the quality of the response outcome was available as early as 80-110 ms before peak force was reached. Interestingly, both the response-locked and TTP-locked analyses (Figure 9C, 10C) showed that the classifier also identified differences between the two error conditions (*too high* vs. *too low*). Furthermore, the absence of any significant time windows for the *correct* vs. *incorrect* analysis (Figure 11) suggested that the classifier was not able to pick up information regarding a more general error-related processes from the brain activity. Added together, these results showed that instead of accuracy-related evidence of the response, information regarding force magnitude was presumably decoded.

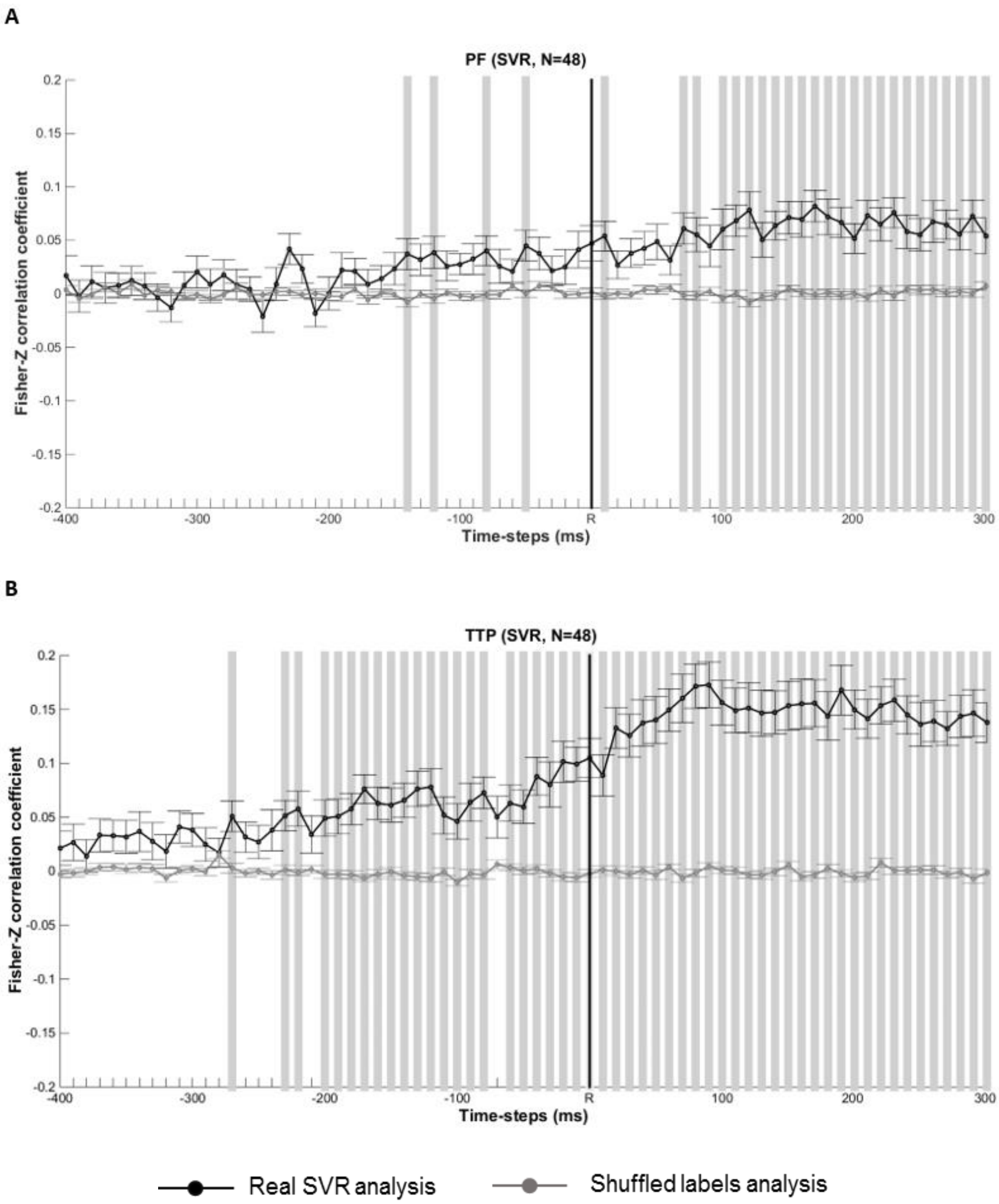


* TTP = time point where Participants reached Peak Force

Figure 11. Classification accuracies of group classification and permutation test result using multivariate pattern analysis (Support Vector Classification, SVC) for correct vs. incorrect (too high & too low) classification; corrected for multiple comparisons (N = 48;

$p < 0.05$; error bars indicate standard errors of the mean; no analysis time window was significant).

Since the classification results suggested that the ERP activity patterns might reflect information regarding force magnitude instead of *correctness*, a further investigation to see *when* specific force parameters of the actually executed force could be predicted was conducted. To address this point, SVR was used to predict the single-trial PF (single-trial TTP) from the ERP patterns. The trials were not separated by conditions any longer but were included in the same analysis. The first SVR analysis (Figure 12A) showed that the PF could be predicted from 140 ms before the response onset up to the point in time of the average TTP. As TTP and PF reflected different aspects of a response (across trials, the mean correlation was small; $r = 0.19$), a further attempt to predict TTP was made in an independent SVR analysis (Figure 12B). These findings showed that TTP, which reflects the temporal dynamics of a force pulse, could be predicted starting from around 270 ms before the response up to 300 ms after the response (remarkably, during the entire time interval after response onset), pointing towards the fact that not only the force magnitude but also the timing of the force pulse was processed (prepared) clearly before the response onset and approximately 100 ms earlier compared to PF.



*R = Response onset

Figure 12. Fisher Z-correlation coefficients and permutation test results using multivariate pattern analysis (Support vector regression, SVR, A, B) for (A) force prediction (PF), and (B) time to peak prediction (TTP); corrected for multiple comparisons (N = 48; $p < 0.05$; error bars indicate standard errors of the mean; grey time windows indicate significant correlation).

2.1.4 Discussion

Until recently, not much information was available regarding the processes involved in action monitoring during tasks with continuous response parameters, such as force production. The first study aimed to fill this gap in the literature and investigated action monitoring during a simple *force execution* using a force production task with one target force range. Three error-related ERP components were assessed, the Ne/ERN/CRN, Pe, and FRN, which are known to be well-established indicators of various error processing mechanisms. Furthermore, a novel multivariate approach (MVPA) was used to investigate the time course of force execution monitoring, *pre* and *post* response. Although clear error-specific variations from the ERP components were not able to be identified, the MVPA results provide first evidences that response related parameters (TTP, PF) are decodable from the brain activity clearly before response was initiated.

2.1.4.1 Force Execution Monitoring

From previous work, it was unclear whether the medial-frontal negative ERP components (i.e., Ne/ERN and CRN) were sensitive to errors committed during force execution, because of some confounding factors in these studies (Armbrecht et al., 2012, 2013; de Bruijn et al., 2003). An attempt to answer this question was made by implementing a *pure* force execution task in this study, therefore eliminating *force selection* process (and therefore potential confounding factors that comes with it). However, no clear error-specific variation between Ne/ERN (*too high*, *too low*) and CRN was observed. The components' sensitivity to the force ranges was also not reflected in the present data, despite the fact that others have reported higher peak amplitudes at the higher force ranges (Armbrecht et al., 2013; de Bruijn et al. 2003). To avoid the amplitude-reducing effects caused by high variation in TTP (Armbrecht et al., 2013), the participants were trained to produce a brief force pulse but were not given a strict TTP limit (and a feedback indicating TTP) in order to focus on correct peak force production. Unfortunately, variability in the TTP was still quite high, which might have affected the components and might explain the unclear findings for Ne/ERN and CRN.

Clear Pe variations were observed between force ranges. Interestingly, the Pe was observable not only in the two erroneous conditions but also in the correct condition. According to the literature, the Pe was often discussed to reflect conscious evaluation of a committed error (see Nieuwenhuis et al., 2001) as well as error evidence accumulation (Steinhauser & Yeung, 2010). For instance, in studies investigating left-right hand errors, a clear Pe was observed in error trials but not in correct trials (see Nieuwenhuis et al., 2001). The existence of Pc in the *correct* condition might reflect the participants' uncertainty when distinguishing a correct from an erroneous response, particularly in trials close to the lower threshold of the correct force range. The higher Pe in the *too high* condition (compared to the *too low* condition) could be resulted from more error evidence accumulation during a shorter time period (Steinhauser & Yeung, 2012) — hence, in the erroneous conditions, participants were more aware of errors made after a stronger key press than a weaker key press.

Another interesting finding is, the FRN amplitude was significantly more negative in the *too low* condition compared to the *correct* condition. However, no significant difference was observed when comparing the FRN in the *too low* condition to the *too high* condition, nor when comparing the *correct* condition to the *too high* condition. According to the *first indicator hypothesis* brought up by Holroyd & Coles (2002), the FRN amplitude should be higher in conditions in which the error was not detected at the time of responding. These results are complementary to the Pe results, which exhibited significant differences only between the *correct* and *too high* conditions, suggesting that these differences were immediately detected, while the other error type: the *too low* errors, was not. Consequently, participants needed feedback to become aware of error responses committed in the *too low* force production range, which would lead to higher FRN amplitudes for the *too low* condition (compared to *correct* condition), but not the *too high* condition. The lack of significant difference in the FRN between the two error conditions, however, provides a challenge to this interpretation. An alternative theory called *the expectancy violation hypothesis* (Hajcak et al., 2005) might explain this specific pattern of results. Previous studies have exhibited that the FRN was larger in conditions with a lower probability of a negative outcome, because a negative outcome was less expected (Hajcak et al., 2007; Holroyd & Krigolson, 2007). If the higher Pe in the *too high* condition reflected successful error detection, the participants should have already expected a feedback showing that an error response was committed. Consequently, there was no violation of expectation in this condition, and therefore a smaller FRN would be observed compared to the *too low* condition, in which the error feedback was particularly surprising, because the unsuccessful error detection in this condition (as suggested by the Pe results). In other words, receiving error feedback in the *too low* condition could have represented expectancy violation, which resulted in a comparably larger FRN.

2.1.4.2 Decoding of Force Production Specific Information from the ERP signals

Previous literature has exhibited that upcoming errors were reflected in the ERP up to 90 ms before an overt response (Bode & Stahl, 2014). Therefore, a set of analyses using a more sensitive method – MVPA – was conducted, to see *if* (furthermore, *when*) information regarding any of the response-related parameters contributing to the quality of a single trial response (i.e., correctness, error types, single trial PF and TTP) was actually able to be decoded from the ERP signal.

By the means of classification analyses, correct trials were able to be distinguished from the two incorrect conditions quite reliably after response onset. The response-locked as well as the TTP-locked classification results for *correct* (target force) vs. *too low* (error) and *correct* (target force) vs. *too high* (error) could be explained by: (1) a variation of force magnitude (i.e., *too low* < *correct* (target force) < *too high*); (2) by error monitoring activity (*correct* vs. *incorrect*); or by (3) a combination of both. However, the classifiers were also able to distinguish the two incorrect conditions from each other. If the ERP patterns solely reflected *correctness* of a response, the classifier should not have been able to identify differences between the two incorrect conditions. Furthermore, this was supported by an additional classification analysis, where two erroneous conditions (*too low* and *too high*) were combined into one 'incorrect' condition, and a decoding attempt of *correct* vs. *incorrect* condition was conducted. The result obtained from conducting this analysis confirmed that the classifier was not able to predict general error-related process from the ERP pattern. Considering both findings, it can be concluded that instead of merely error specific information, force magnitude information was decoded by the classifier. What is more, if this was the case, the additional *correct* vs. *incorrect* classification would give the expected result, since both "incorrect" patterns (for the *too high* and *too low* conditions) would be marked differently from each other, and the correct pattern would then lie in-between them, reducing the classifier's ability to reliably distinguish differences between conditions. It is also remarkable that the significant classification results were distributed consistently starting from the time of responding until 300 ms after response onset, and therefore less likely to only mirror differences in the ERP components described above. While on average most accuracies did not significantly differ from chance, accuracies were enhanced in some time windows (even before response onset). This could point to some trials during which the final PF was determined very early, which in turn was reflected already in the brain activity patterns. The cognitive origins of these patterns are rather diffuse. However, it remains difficult to clearly relate these results to specific cognitive processes based on these analyses alone. Another noteworthy thing is, the methods used in this study were aimed to estimate a lower bound which is conservative and can be trusted (i.e. not to be a *false positive* result). There are a variety of ways to optimize classifiers' accuracy (e.g. modification of the analysis time window width, steps reduction in the cross-validation process, feature elimination, and many other ways). A decision to sacrifice absolute accuracy was made in favour of statistical rigour, since absolute accuracy could falsely imply that it is a measure of how much of the process can be predicted from the brain activity pattern. Instead of optimizing accuracy level, it is more important to ensure that the obtained accuracy values were truly above chance. Note that researches in the field adapted a similar approach, with accuracy rarely exceeding 60% (e.g., Fahrenfort, Grubert, Olivers, & Eimer, 2017; Hogendoorn & Burkitt, 2018).

The following SVR analyses revealed that the single-trial PF could already be decoded from the ERP pattern at least 140 ms before the response onset. This suggests that the information regarding the response outcome was clearly available before response was initiated in the periphery (i.e. when the minimum force of 50 cN required to be considered as a button press was exceeded). Similar to this, a previous study incorporating force sensitive keys has shown that the response quality (correctness) of simple decisions (i.e., erroneous responses during flanker task) could be decoded around 100 ms before an overt response was registered (see Bode & Stahl, 2014). However, this previous study did not attempt to make prediction of the actual response force, and differences in force ranges were not considered as a component of the task either. Taken together, results of the current study suggest that the brain seems to process information that is indicative of subsequent response errors long before response initiation. However, this does not necessarily mean that this information is available to use for immediate response monitoring, as it could simply reflect neural patterns which are associated with incorrect response planning.

The second SVR analysis showed that the single-trial TTP was decodable during the entire period after response onset even more consistently than the PF. Furthermore, this result pointed out that the temporal dynamics of force production —reflected by TTP— is likely to be affected by processes even earlier than the maximum of the force itself, since it was decodable already 270 ms before response onset (the PF was – in comparison to

TTP – only decodable starting from 140 ms before response onset). It was suggested by Ulrich et al. (1995) that TTP and PF are not fully independent parameters, but more than just two sides of the same coin. This is supported with one of the findings in this study; that both parameters were only marginally correlated. The *Parallel Force Unit Model* (Ulrich & Wing, 1991) provides a reasonable explanation for this partial independence: when participants produced force in a specific time, they would need to adjust two response parameters: the duration of force unit activation (which is reflected in the TTP) and the number of recruited force units (i.e., force-producing motor units - reflected in the PF).

Combined findings from both SVR analyses (for PF and TTP) provided important evidence that the timing and magnitude of a simple one-dimensional force pulse were planned already before response execution. However, since a force pulse is regarded as a ballistic process (Desmedt, 1982; Cordo, 1987), with some controllable aspects, like by the modification of the number of the involved force units (Ulrich & Wing, 1991), an early definition of the response parameters in the motor program is likely to be an efficient strategy for a fast force production, and therefore might serve as information for an error detection process. However, this early force magnitude monitoring process seemed to have happened before the process of determining the *correctness* of a certain response (if they were not completely independent), in relation to 'the default force' required for a response to be deemed correct. In consequence, even though the brain seemed to be fast enough in term of planning the response parameters, the *correctness* aspect of the response itself was not yet foreshadowed in the neural activities.

2.1.5 Limitations

The first set of analyses in this study was conducted to see if the classical ERP components reflected error monitoring, force monitoring, or a mixture of both processes. However, the ERP findings showed no clear differentiation between error processing activity from force-time related activity, which is presumably related to the nature of the force task used in the study, as the information about the response quality (*correctness*) is naturally contained in the force magnitude and modulated by its temporal dynamic. The multivariate analyses results showed that the magnitude of response force – instead of *correctness* could be decoded from the brain. However, it is important to consider that there is still a possibility that error specific information was in fact encoded in the brain activity, and the method used in the first study might not have been sensitive enough to pick up this information. Another limitation is that although the trial numbers for each condition for each classification analysis was balanced (which always resulted in equal number of trials for both conditions in all training and test sets), there is always a risk that differences in variance between conditions could bias the classifier. However, there was no evidence for differences in variance in the higher-dimensional space in the current data's patterns, and the SVR analyses did not include different conditions, which makes it unlikely that any systematic biases have occurred.

2.2 Study 2

2.2.1 Objective of the Study

The second study was designed to investigate the course of two subsequent simple force production tasks. During subsequent tasks, depending on the length of the interval between the two tasks, response execution for the second task could be impaired (i.e., resulting in a prolonged response time for the second task). Such phenomenon has been widely known as a 'PRP effect', and has been found in studies using subsequent discrete choice tasks, such as flanker tasks. In this second study, a paradigm was specifically designed to investigate if response execution during two subsequent force production tasks was modulated by a similar PRP-like effect. Note that in a simple choice-response task, a 'response' is considered discrete, since it does not require further monitoring and adjustment after initiation (to illustrate, a decision process is completed once the brain has decided for one response option as the 'correct' answer, and a response is therefore completed after the chosen response is given– for example, when a button corresponding for the chosen option is pressed). However, as opposed to a simple *n*-choice response task, force production task is more complex in its execution. For instance, the accuracy of a force response is defined as a range on the force continuum, which means, the response planning process (i.e., force programming) starts before responding (as exhibited in the first study's result) and the force monitoring process does not end at the time when a response was initiated. Instead, this monitoring process goes on at least until the PF is reached, during which the brain has to continuously control the amount of Force Units produced to ensure that the force execution is going as planned. Consequently, instead of varying interval between the two subsequent stimuli (i.e., *Stimulus Onset Asynchrony* / SOA) which is commonly used in PRP studies using discrete choice tasks, for the investigation of force production the time interval between the first response and the onset of the second stimulus, the so called *Response-Stimulus Interval* (RSI), is more relevant in the present study.

The first aim of this study was to replicate the PRP effect, that is the delay of the response time for the second task, which usually arises during two consecutive discrete choice tasks, in an 'adapted' PRP paradigm designated for two consecutive simple force tasks. However, beyond replicating this effect, this study was further aimed to investigate how different *Response-Stimulus Interval* (RSI) affect the cognitive processes involved in force production, not only on a behavioral level (i.e., response time and response force parameters), but also on a neural level (i.e., force monitoring-related ERP components). To investigate PRP effect on a neural level, Ne/ERN (CRN for the correct condition) was tested, to see how the different RSIs modulated the monitoring activity of the second force task. If the PRP effect happened to modulate error monitoring activity of the second task, Ne/ERN (CRN) should be lower in the short RSI condition, and gradually increased in the longer RSI conditions. The Pe (Pc) component was also investigated, as it reflects accumulation of error evidence and thus might be an interesting measure in term of force-related error detection. Furthermore, Pc (Pe in the correct condition) is also known to be linked to the level of uncertainty while performing a task, which could be an indicator that one or more parts of the information processing stages were interrupted. Thus, investigating Ne/ERN (CRN) and Pe(Pc) could provide evidence on how the information processing stages in the current paradigm was modulated by RSI manipulation, and therefore allows for better understanding on how two subsequent force production tasks unfold – depending on the length of 'pause' between the two tasks.

The FRN and Feedback P3, which were related to the external feedback-processing, were also investigated in this study. In force production tasks, external feedback provides important information such as the precise response force produced during a trial - which is difficult to obtain only through internal action monitoring, especially considering that the participants used in this study were not motor experts. Thus, not only that (external) feedback is needed to 'confirm' that a correct response was produced, it is also necessary for response adaptation process (i.e., planning the motor program for the next trial), after under/over-producing a response force. Furthermore, one could look into these feedback-related components to detect any discrepancy between the *expected* and *received* outcomes of both tasks, which could provide a clearer picture regarding the participants' behaviour while processing external feedback, thus allowing a better understanding regarding the flow of the information during the execution of both tasks.

2.2.2 Method

2.2.2.1 Participants

Forty-eight participants (39 females), all students of the University of Cologne, participated in this study. Due to insufficient number of errors (less than 6 trials per condition) and technical artefacts, data from 8 participants were excluded. The analyses reported here were based on data from the remaining 40 participants (34 females, age range: 18 to 42 years; mean ± SD: 24.47 ± 5.47 years). All participants had normal or corrected-to-normal vision, and were right handed. None of the participants reported to be ambidextrous. Informed consent was obtained from all participants, and they were rewarded with course credit for their participation.

2.2.2.2 Apparatus

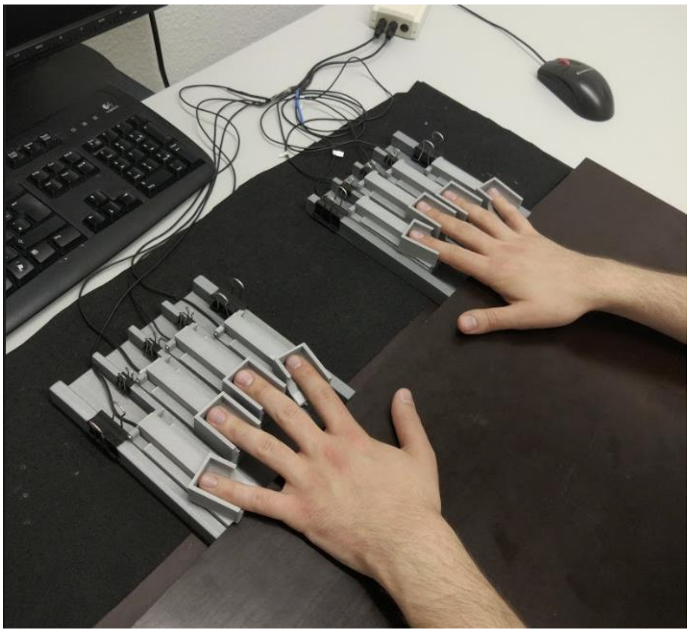


Figure 13. Experimental Setup: Force keys were embedded in an adjustable custom made 3D-printed plastic structure to measure response force; participant pressed the free end of the key to produce an analogue electrical signal corresponding to the produced force.

A different set of force sensitive keys (Figure 13) was used to record the behavioural data. In each key, a force sensor (FCC221-0010-L, DigiKey MSP6948-ND) was installed. These force sensitive keys were embedded in a 3D-printed plastic structure whose positions were adjustable to the accommodate the varying size of the participants' hands and fingers. Note that although these keys can record force response from eight fingers, only two keys were used; one for the left index finger (i.e., for the first button press) and one for the right index finger (i.e., for the second button press). A force response was registered as soon as the keys were pressed with a force of 45 cN or more. The analogous signal was digitized by using a VarioLab AD converter (Becker-Meditec) at a sampling rate of 1024 Hz with a resolution of 16 bits. A photo sensor was additionally attached to this converter, which enabled the detection of a stimulus onset (on the screen) through a change of brightness. This mechanism allows for events assessments (i.e., stimuli, responses, feedbacks) in a real time without noticeable temporal delay. The participants were instructed to lay their arms on a board placed between them and the computer screen, and an adjustable chin rest was installed around 85 cm from the screen to maintain a stable posture. Prior to the experiment, participants were required to adjust the force keys and the provided chin rest to ensure a comfortable and a stable posture throughout the experiment. Before the experiment started, the keys were calibrated to the individual finger weight.

2.2.2.3 Maximum Voluntary Force

At the beginning of the first and the second training session (see below for details), participant's *maximum voluntary force* (MVF) for the left and right index finger were separately assessed, to determine the individual force ranges for each finger (defined in % MVF; see below), as MVF varies across participants. The participants were instructed to press the force key with their left or right index finger as hard as possible without moving the forearm, and repeated this procedure seven times in a row. A start signal ("JETZT DRÜCKEN" German for 'push now') initiated each of the seven key presses. The individual MVF was calculated by averaging the PF of the last four key presses. Three force ranges were then defined relative to the individuals' MVF (i.e., for the right index finger, *target range*: 43–58% MVF; *too high*: > 58% MVF; *too low* < 43% MVF; for the left index finger, *target range*: 37–52% MVF; *too high*: > 52% MVF; *too low* < 37% MVF). The decision to slightly differ the target ranges of the two fingers was made after taking into accounts the results of a previously conducted pilot study (with 15 participants), and was aimed to control for a similar difficulty level between the two tasks, as a previous study (see Lien et al., 2003) have shown that PRP effect can be nullified by making the first task easier.

2.2.2.4 Experimental Task

Participants were tested individually in a 50-minute experimental session, preceded by a 20-minute training session (see Figure 14) which was divided into three parts. The first part of the training session (50 trials) was aimed to train the participant to reach the target force range with their left index finger. Each trial began with a stimulus presentation of two squares on a black background; the left square was colored in white and the right square was a white frame fill in black. The white-colored left square indicated that the participants were required to produce a brisk, isometric force pulse with their left index finger, as soon as this picture appeared on the screen. On each trial, the participants were asked to reach the middle range of their maximum force (represented by the green area in the middle of the *force ruler*) as quick as possible. In addition, they were required to reach the peak force within 180 ms after response initiation. If they could not reach the peak force within 180 ms, they would receive an error message 'Taste wurde so lang gedrückt' (German for 'button was pressed too long') and would not be presented with a feedback. To encourage the participant to make a quick response, they would also receive an error message 'zu langsam' (German for 'too slow') when a button press was not initiated within 500 ms after the stimulus onset. They received a feedback when they responded quick (RT < 500 ms) and fulfilled the TTP requirement of 180 ms. When the participants made at least 30% correct responses in the first training part, they were allowed to proceed to the second part. The second part of the training session was exactly the same as the first; the presented stimulus consisted of two squares, except that the right square was filled in white (while the left white square frame was filled in black), and the participants were required to produce a brisk, isometric force pulse with their right index finger.

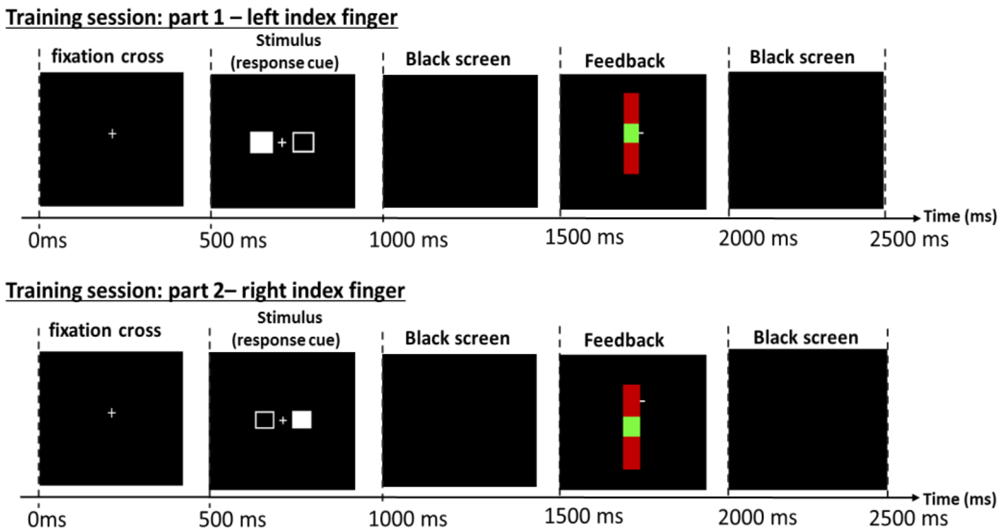


Figure 14. Separate training session for the left and right index finger

When the participants made at least 30% correct responses in the second training part, they were allowed to proceed to the third part (Figure 15), which was identical to the actual experiment. In this last part of the training session, the participants had to respond to two subsequent stimuli. They were required to press the force key with their left index finger when shown the first stimulus (i.e., two squares with the left square colored in white) and then press the force key with their right index finger when shown the second stimulus (i.e., two squares with the right square colored in white). The time interval (RSI) between the onset of the first response and the second stimulus onset was varied without the knowledge of the participants (i.e. lag 1: 550 ms, lag 2: 700 ms, lag 3:1100 ms; lag was varied within block). After making two subsequent button presses (first with the left index finger, followed by the right index finger), two individual feedbacks were displayed one after another. The first feedback was presented for 500 ms to indicate whether the correct target range for the first response (by the left index finger) was reached, and the second feedback was also presented for 500 ms to indicate whether the correct target range for the second response (by the right index finger) was reached. The green area in the middle of the force ruler is a representation of the *correct* force range. The upper red area corresponded to the *too high* condition, and the bottom red area corresponded to the *too low* condition. The produced force level in the current trial was indicated by a white cursor. The actual experiment (Figure 15) was exactly the same as the third training part, except that it consisted of 7 blocks with 50 trials in each block. Between each block, the participants were forced to make a brief 3-minute rest.

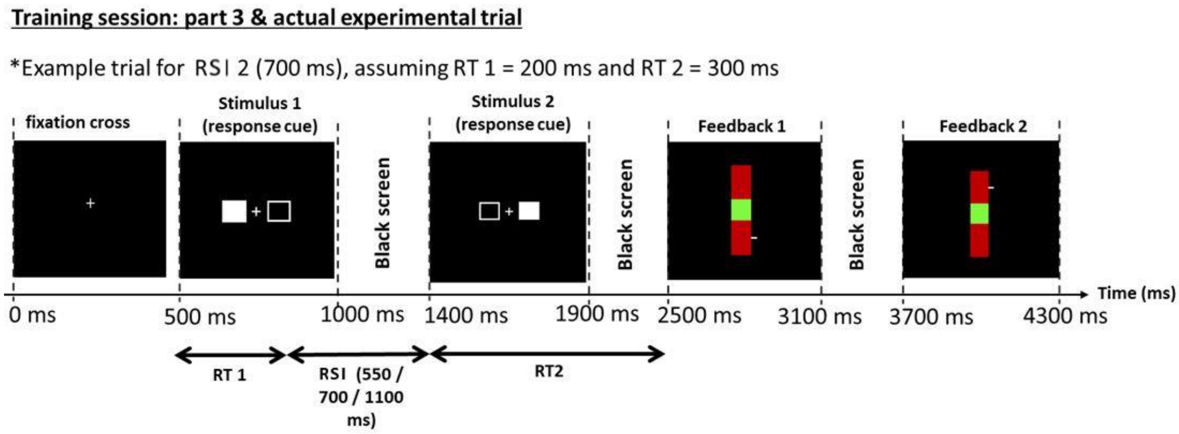


Figure 15. Training session part 3 and actual experimental trial

2.2.2.5 Data Acquisition

Behavioral Data

RT was defined as the first time point when the participant's response force exceeded 45 cN (i.e. response onset), measured from the stimulus onset. PF was defined as the maximum of the force (in cN) of a single trial force pulse. TTP was defined as the time point (measured from response onset) at which the peak force was reached. Frequency of correct responses, frequency of *too high* responses, frequency of *too low* responses, mean RT, mean PF, and mean TTP were determined separately for each button press in each condition and for each participant.

Electrophysiological Data

EEG data were obtained from 61 scalp electrode sites according to the standard international 10–20 system (Jasper, 1958). The active Ag/AgCl electrodes (actiCAP, Brain Products) were referenced against the left (i.e., active reference) and right mastoid (i.e., passive reference). The electrooculogram (EOG) signals were recorded using passive bipolar Ag/AgCl electrodes (ExG-Amplifier, Brain Products). The vertical EOG was recorded above and below the left eye, and the horizontal EOG was recorded at the left and right temples. The EEG was continuously recorded at a sampling rate of 500 Hz using BrainAmp DC (Brain Products).

The electrophysiological data for each button press were time-locked to (a) the response onset; and (b) the feedback onset. Data were epoched ranging from 100 ms before until 500 ms after response (or feedback) onset. Baseline correction was performed with the period of 100 ms before response onset (or feedback onset, for the feedback-locked data). An ocular correction algorithm was applied in order to reduce the impact of eye movements (Gratton, Coles, & Donchin, 1993), followed by a second baseline correction. Afterwards, an artefact-rejection procedure was carried out to eliminate contaminated trials exceeding maximum/minimum amplitudes of ± 100 μV. The remaining trials were averaged. A current source density (CSD) analysis was then performed on the averaged ERP waveforms. This analysis accounted for the curvature of the head using a spline algorithm (Perrin et al., 1989), and was performed to reduce the effect of neighbouring currents. The CSD signals (order of splines = 4; lambda = 10^{-5}; maximal degree of Legendre polynomials = 10) were computed for each electrode site by taking the second derivative of the distribution of the voltage over the scalp. Lastly, the ERP components (Ne/ERN, CRN, Pe/Pc for correct trials, FRN, and Feedback P3; both peak amplitudes and mean amplitudes as the standard indicator for the area under the curve in the defined time ranges – hereafter referred to as 'area') were determined separately from the individual mean CSD-ERP waveforms at the electrode sites FCz and Cz. Ne/ERN (CRN) was defined as the most negative peak in a time window ranging from 0 to 180 ms after response onset at electrode site FCz. Pe (Pc) was defined as the most positive peak in a time window ranging from 150 to 300 ms after response onset at electrode site Cz. FRN was defined as the most negative peak in a time window ranging from 150 to 250 ms after feedback onset. Finally, the Feedback P3 was defined as the most positive peak in a time window ranging from 260 to 460 ms after feedback onset. Both FRN and Feedback P3 were investigated at the electrode site FCz.

2.2.2.6 Statistical Analyses

RSI Effect on Behavioral Data and Neural Correlates of Force Monitoring (Response –locked Averages)

Several repeated-measures analysis of variance (ANOVAs) were conducted for the within-subject factor Response-Stimulus Interval / RSI (*short*, *medium*, and *long RSI*) for all behavioural measures (RT, PF, TTP) separately for each button press (i.e., left hand and right hand) using SPSS 23. An additional ANOVA with the within-subject factor *error rates* (frequency of errors for the left hand and the right hand) was conducted. Note that, data from all trials and all force conditions (*correct, too high, too low*) were used for the behavioural data. Another set of ANOVAs with the within-subject factor RSI (*short, medium, and long RSI*) were further performed for the respective ERP components' peak amplitudes (Ne/ERN and CRN at the electrode side FCz, Pe/Pc at the electrode side Cz) and the components' area measures, separately for each button press (left hand and right hand) and each force condition (*correct, too low*). The *too high* condition was excluded from the ERP analyses due to insufficient number of trials (less than 6 trials for this condition). Additionally, to investigate *force range* effect, several two-way ANOVAs with the within-subject factors *force range* (*correct, too low*) and RSI (*short, medium*, and *long RSI*) were performed separately for the Ne/ERN or CRN and Pe/Pc peak amplitudes and area measures. Note that these additional analyses were aimed to see variations between the force conditions (*correct, too low*), thus interaction effect and RSI were not relevant and therefore were not further investigated. Significant ANOVA results were followed up using Bonferroni adjusted post-hoc tests. Level of significance were adjusted using Geisser and Greenhouse (1958) correction in case the sphericity assumption was violated. Effect sizes are reported in terms of partial eta² (η_p^2).

Force Effect on Feedback Processing (Feedback-locked Averages)

Several two-way ANOVAs with the within-subject factors *force range* (*correct, too low*) and RSI (*short, medium,* and *long RSI*) were performed separately for the FRN peak amplitude and the Feedback P3 peak amplitude (and areas) for each button press, at the electrode side FCz. Note that factors other than *force range* (i.e., RSI and interaction effect) were not of interests in the present study, thus were not further investigated. Significant ANOVA results for force were followed up using Bonferroni adjusted post-hoc tests. Level of significance were adjusted using Geisser and Greenhouse (1958) correction in case the sphericity assumption was violated. Effect sizes are reported in terms of partial eta² (η_p^2).

2.2.3 Results

2.2.3.1 Behavioral Data

The first set of analyses was conducted to see whether the RSI affected RT, PF, and TTP of the second button press (see Figure 16). For these analyses, data from all conditions (*correct*, *too high*, and *too low*) were used. The same analyses were also conducted for the same parameters (RT, PF, and TTP) of the first button press (left index-finger response) as a sanity check (i.e., to confirm that the PRP effect was exclusive to the second button press). The respective ANOVA for the second button press showed that RSI had a significant effect on Response Time, $F(2,38) = 1585.853$, $p < 0.001$, $\eta_p^2 = 0.987$. The slowest response was observed in the short RSI condition (475.66 ± 6.62 ms), followed by the medium RSI condition (309.46 ± 5.97 ms), and the long RSI condition (168.48 ± 4.56 ms). Follow-up post-hoc tests confirmed significant differences between the short and long RSI condition ($p < 0.001$), the short and medium RSI condition ($p < 0.001$), as well as the medium and long RSI condition ($p < 0.001$).

In term of PF, ANOVA results showed significant RSI effect on the second button press, $F(2,38) = 22.403$, $p < 0.001$, $\eta_p^2 = 0.541$. The highest force was observed in the short RSI condition (990.91 ± 31.50 cN), followed by the medium RSI condition (975.49 ± 31.37 cN), and the long RSI condition (944.18 ± 31.28 cN). Follow-up post hoc tests confirmed significant differences between the short and medium RSI condition ($p = 0.002$), the medium and long RSI condition ($p < 0.001$), as well as the short and long RSI condition ($p < 0.001$).

Significant RSI effect was also observed on TTP for the second button press, $F(2,38) = 10.913$, $p < 0.001$, $\eta_p^2 = 0.365$. The longest TTP was observed in the medium RSI condition (169.54 ± 5.89 ms), followed by the short RSI condition (167.75 ± 6.22 ms), and the long RSI condition (157.47 ± 5.06 ms). Post-hoc t-test result for TTP confirmed significant differences between the short and long RSI condition ($p = 0.004$) and between the medium and long RSI condition ($p < 0.001$).

Lastly, RSI effect on error commission rates was investigated (see Figure 17). No significant RSI effect was observed on error commission rates for the second button press, $F(2,38) = 0.305$, $p = 0.739$, $\eta_p^2 = 0.016$. No significant RSI effect was observed for the first button press for RT ($p = 0.555$), response force ($p = 0.098$), TTP ($p = 0.124$), and error commission rates ($p = 0.282$)

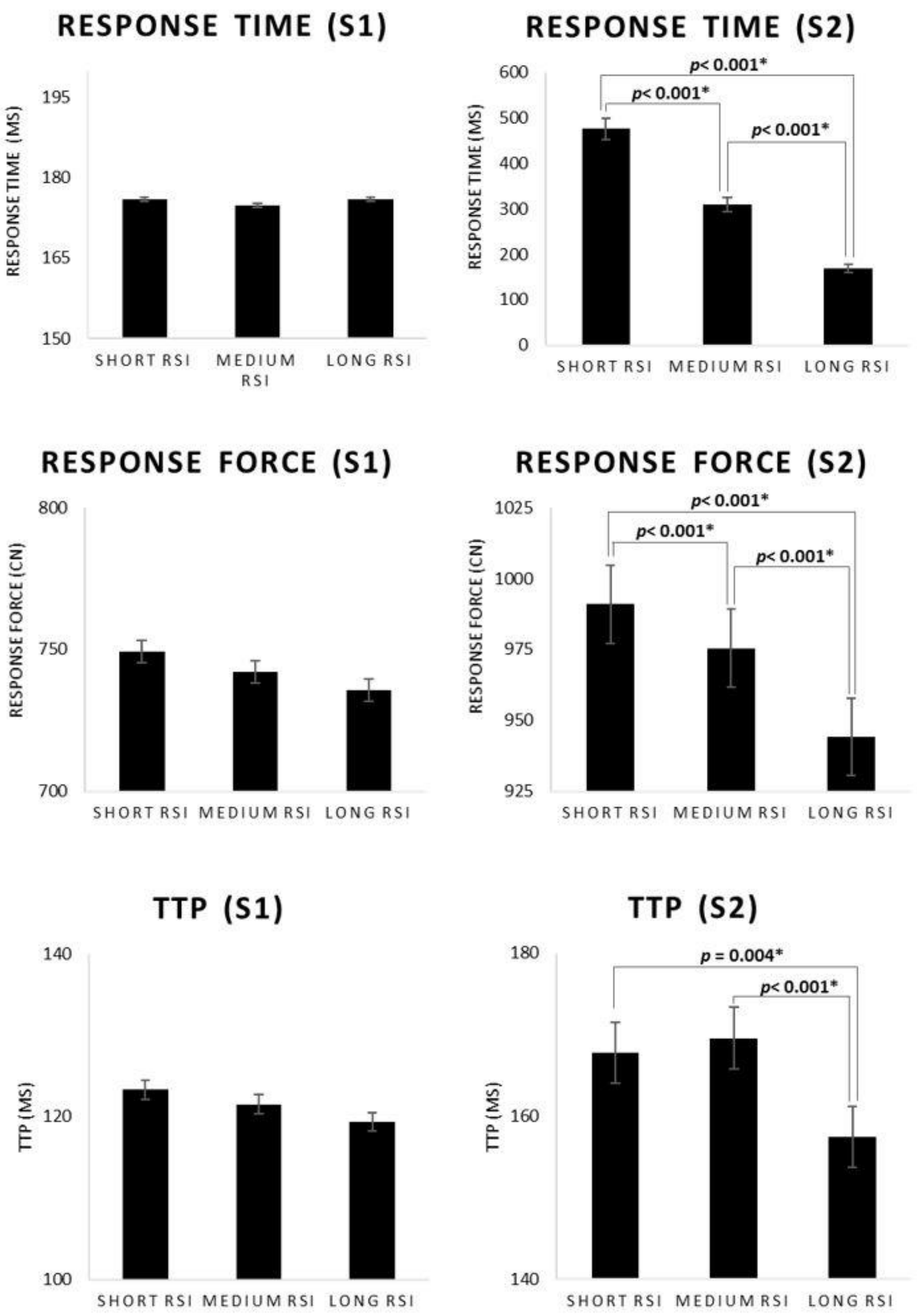
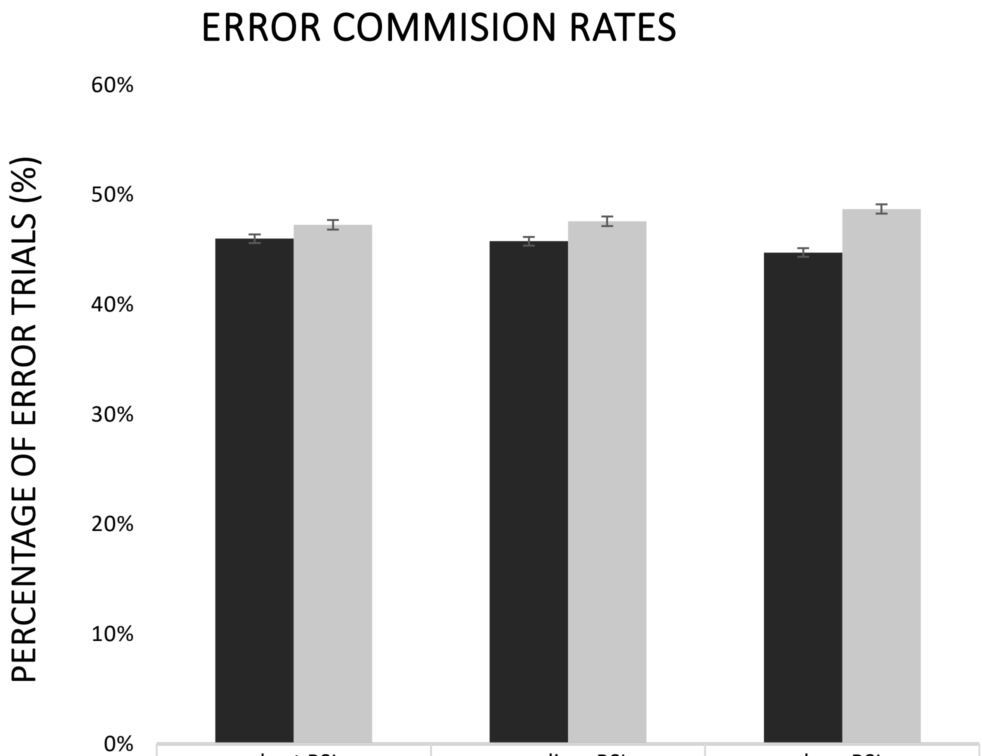


Figure 16. RSI effect on behavioral data for the first button press (S1 – left index finger) and the second button press (S2 – right index finger). Error bars represent standard errors of means. Significant differences between conditions are marked with asterisks.



	short RSI	medium RSI	long RSI
S2 (right-index finger)	46%	46%	45%
S1 (left-index finger)	47%	48%	49%

Figure 17. RSI effect on error commission rates for the first button press (S1 – left index finger) and the second button press (S2 – right index finger). Error bars represent standard errors of means.

2.2.3.2 Electrophysiological Data

RSI Effect on Response-locked Averages

Multiple repeated measures ANOVAs were conducted separately for each force condition (conditions: *too low*, *correct*), to test the RSI effect on the peak amplitude and area for Ne/ERN and Pe after the second button press. In the *too low* condition, significant RSI effect was observed in the Ne/ERN area for the second button press, $F(2,38) = 4.397$, $p = 0.020$, $\eta_p^2 = 0.101$ at FCz (see Figure 18). Post-hoc test results showed significant differences for Ne/ERN area between the short RSI (26.29 ± 2.03 μV/cm^2) and the medium RSI condition (31.71 ± 2.40 μV/cm^2, $p = 0.002$). No significant difference was observed between the long RSI condition (29.44 ± 2.19 μV/cm^2) and the other two conditions ($p >$

0.402). No significant RSI effect was observed in the Ne/ERN peak amplitude for the second button press, $F(2,47) = 0.88$, $p = 0.423$, $\eta_p^2 = 0.045$.

In the *correct* condition, significant RSI effect was observed in the CRN area for the second button press, $F(2,38) = 4.940$, $p = 0.012$, $\eta_p^2 = 0.206$ at FCz. Post-hoc results (Figure 19) showed significant differences for CRN area between the short (26.39 ± 1.69 μV/cm^2) and the medium RSI condition (31.65 ± 2.44 μV/cm^2, $p = 0.012$), as well as between the short (26.39 ± 1.69 μV/cm^2) and the long RSI condition (31.42 ± 2.66 μV/cm^2, $p = 0.049$). Highly similar results patterns for main effect and post hoc comparisons were observed in the CRN peak amplitude for the second button press, $F(2,38) = 5.204$, $p = 0.010$, $\eta_p^2 = 0.215$.

As a sanity check, a series of ANOVAs was also conducted for the first button press to see if different RSIs affected Ne/ERN (or CRN) in both – the *too low* condition and the *correct* condition. The ANOVA results for the first button press showed no significant effect of RSI for Ne/ERN peak amplitude ($F(2,38) = 1.198$, $p = 0.313$, $\eta_p^2 = 0.061$) and area ($F(2,38) = 0.045$, $p = 0.956$, $\eta_p^2 = 0.002$) in the *too low condition*, as well as for CRN peak amplitude ($F(2,38) = 0.889$, $p = 0.420$, $\eta_p^2 = 0.045$) and area ($F(2,38) = 1.625$, $p = 0.210$, $\eta_p^2 = 0.079$) in the *correct* condition.

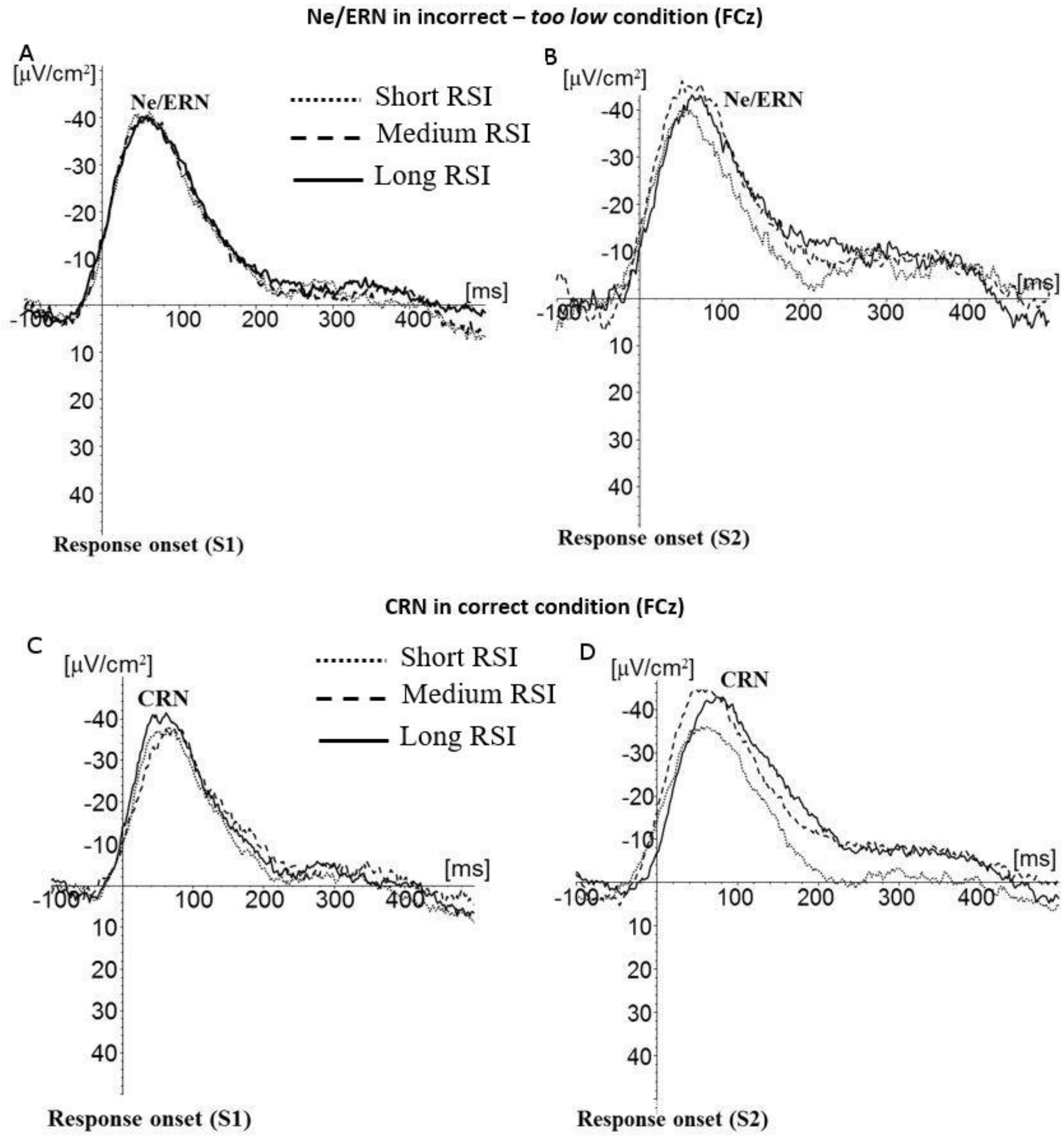


Figure 18. Averaged event-related potentials of the force task for the different RSI conditions (short, medium, and long), time-locked to the response onset, for the first button press (A, C) and the second button press (B, D), in the *too-low* condition (A, B) and the *correct* condition (C, D), at the electrode side FCz, for error/correct-related negativity (Ne/ERN or CRN).

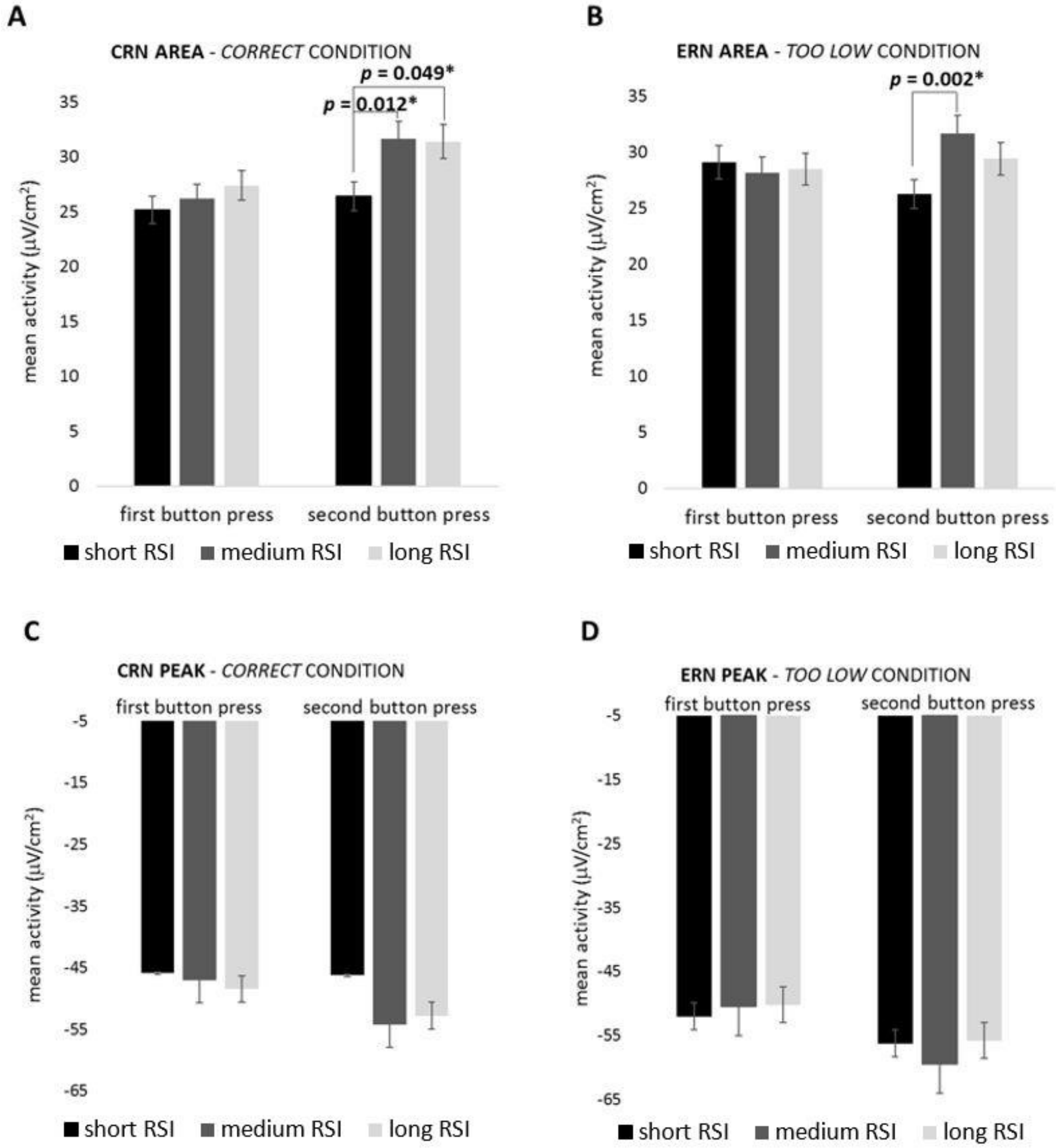


Figure 19. Post hoc results on RSI effect for the Ne/ERN (CRN) areas (A, B) and peak amplitudes (C, D), in the *correct* condition (A, C), and the *too low* condition (B, D), for the first and second button press. Significant differences between conditions are marked with asterisks.

Significant effects of RSI for the second button press were detected for both response-locked Pc peak amplitude, $F(2,38) = 7.111$, $p = 0.023$, $\eta_p^2 = 0.099$, and Pc area, $F(2,38) = 4.369$, $p = 0.020$, $\eta_p^2 = 0.191$, at Cz for the *correct* condition (Figure 20). Post hoc results for Pc peak amplitude (see Figure 20) showed significant difference between the short RSI (20.52 ± 2.65 μV/cm^2) and medium RSI (13.86 ± 2.38 μV/cm^2, $p = 0.006$)

condition, and *marginally significant* difference between the medium RSI (13.86 ± 2.38 μV/cm^2) and the long RSI (19.50 ± 3.06 μV/cm^2, p = 0.073) condition. While for the Pc area, significant difference was observed between the medium (12.46 ± 1.10 μV/cm^2) and long RSI (19.50 ± 3.06 μV/cm^2, p = 0.015) condition. No significant difference was observed between the short RSI condition (12.46 ± 1.10 μV/cm^2) and the other two conditions (p > 0.50). ANOVA results for the Pe peak amplitude for the second button press in the *too low* condition showed no significant RSI effect, although it is noteworthy to mention that the p value almost reached the level of significance, $F(2,38)$ = 2.697, p = 0.081 , η_p^2 = 0.127. Further post hoc result also showed marginally significant difference between Pe peak amplitude between the short RSI (25.18 ± 3.12 μV/cm^2) and the medium RSI condition (18.63 ± 2.53 μV/cm^2, p = 0.072). No difference was observed between the long RSI condition (21.05 ± 3.24 μV/cm^2) and the other two conditions (p > 0.50). No significant RSI effect for the second button press was observed for Pe area, $F(2,38)$ = 1.145, p = 0.329 , η_p^2 = 0.058.

As a sanity check, a series of ANOVAs was also conducted for the first button press to see if different RSIs affected Pe (Pc – in the *correct* condition) in both – the *too low* and the *correct* condition. The ANOVA results for the first button press showed no significant effect of RSI for Pe peak amplitude ($F(2,38)$ = 1.053, p = 0.359 , η_p^2 = 0.055) and area ($F(2,38)$ = 1.219 , p = 0.307 , η_p^2 = 0.062) in the *too low condition*, as well as for Pc peak amplitude ($F(2,38)$ = 1.161, p = 0.324, η_p^2 = 0.058) and area ($F(2,38)$ = 1.799, p = 0.179, η_p^2 = 0.086) in the *correct* condition.

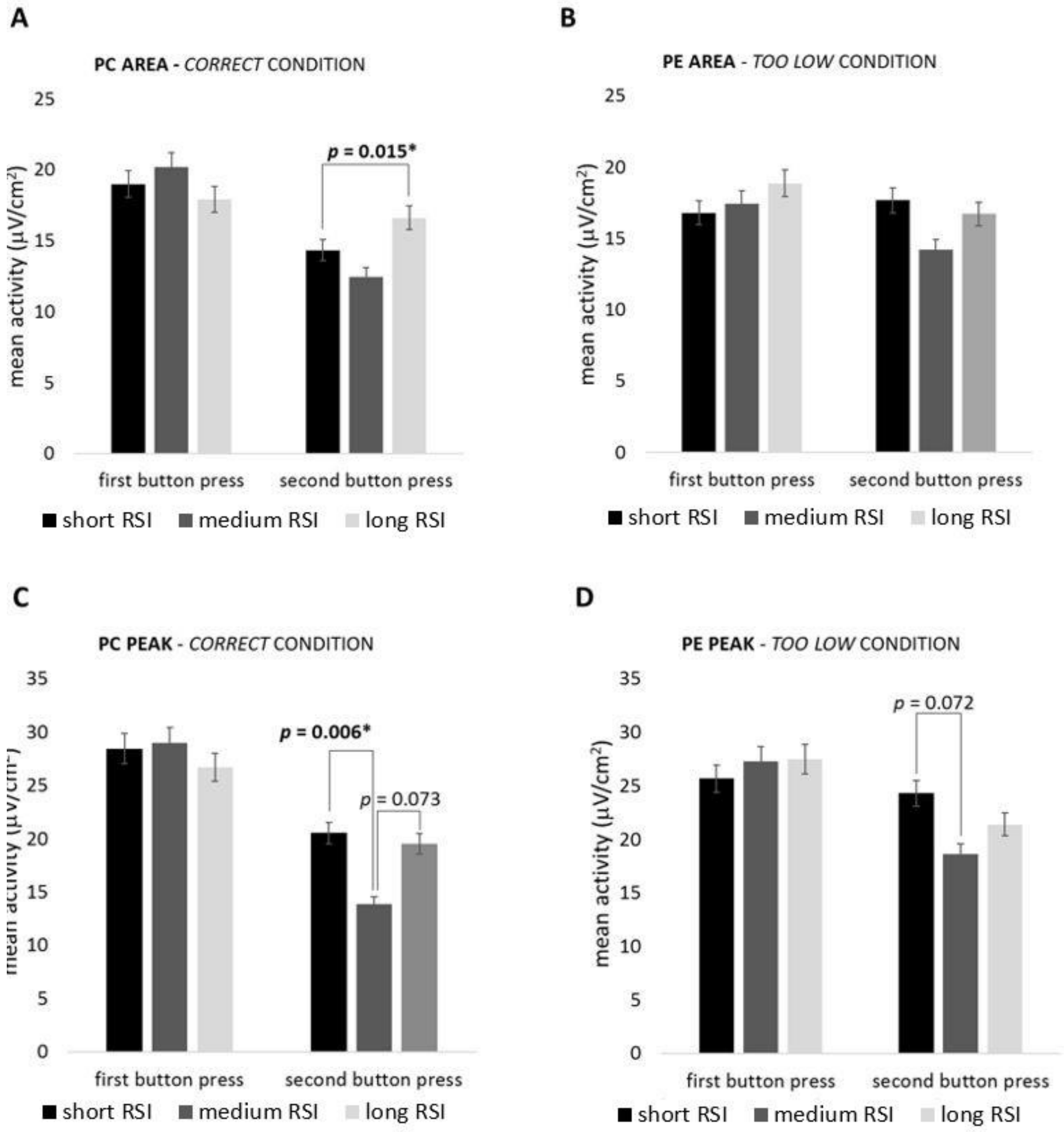


Figure 20. Post hoc results on RSI effect for the Pe(Pc) areas (A, B) and peak amplitudes (C, D), in the *correct* condition (A, C) and the *too low* condition (B, D), for the first and second button press. Significant differences between conditions are marked with asterisks.

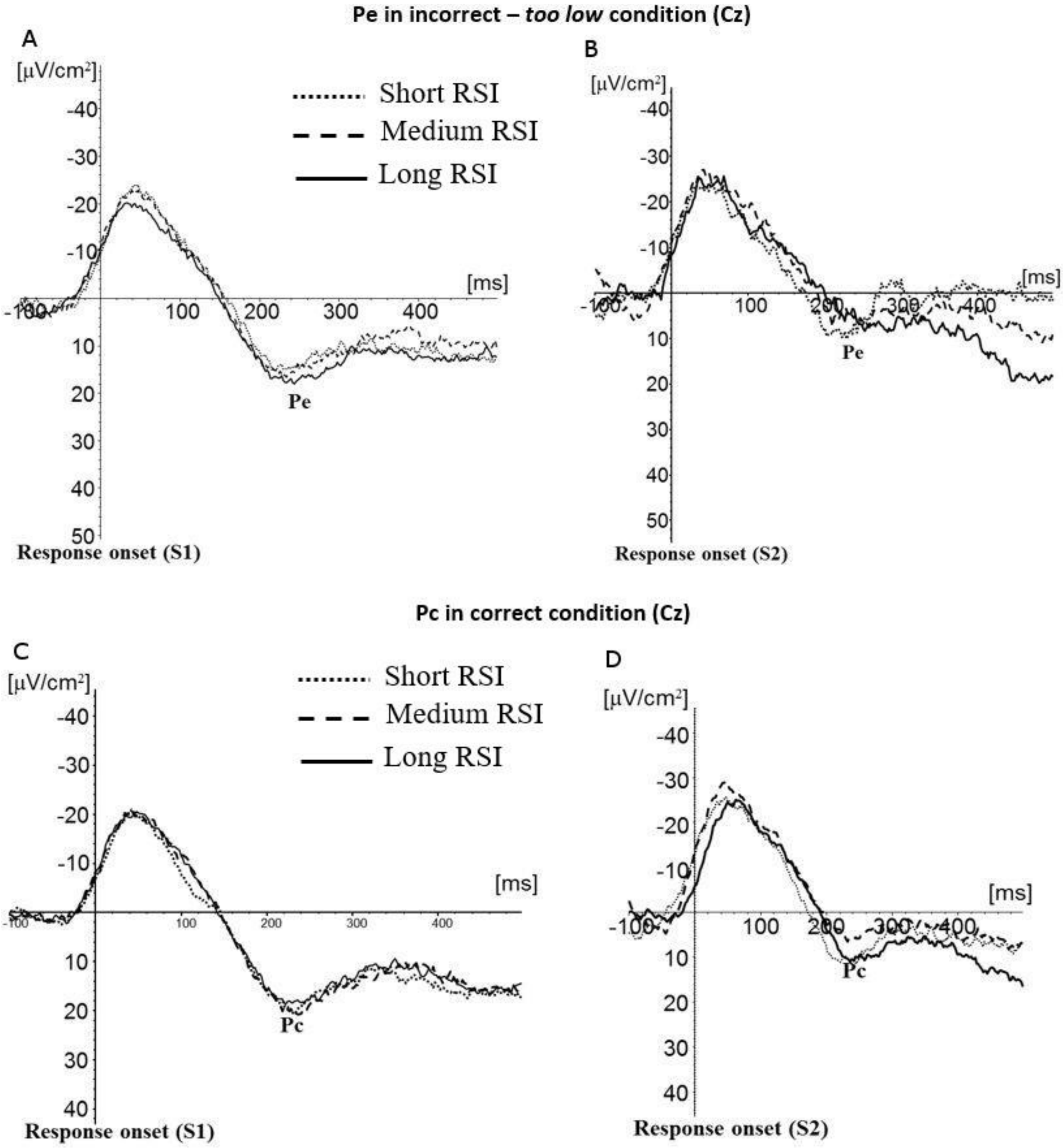


Figure 21. Averaged event-related potentials of the force task for the different RSI conditions (short, medium, and long), time-locked to the response onset, for the first button press (A, C) and the second button press (B, D), in the *too-low* condition (A, B) and the *correct* condition (C, D), at the electrode side Cz, for error/correct positivity (Pe or Pc).

Force Effect on Response-locked Averaged Event-related Potentials

A series of repeated measures ANOVAs were carried out separately for the first and second button press (i.e. left-index finger and right-index finger) to see if there are difference(s) between the two force conditions (i.e. *correct* and *too low* conditions) in term of ERP components (i.e. Ne/ERN and Pe/Pc peak amplitudes and areas) in the time window of interest, specific for each ERP component (see Figure 21). The first series of ANOVA results for the second button press showed marginally significant difference between the Pe peak amplitude (21.62 ± 2.35 $\mu V/cm^2$) and the Pc peak amplitude (18.51 ± 2.30 $\mu V/cm^2$), $F(1,39) = 3.297$, $p = 0.077$, $\eta_p^2 = 0.080$. There was no significant force effect observed for the Pe(Pc) area, $F(1,39) = 0.593$, $p = 0.446$, $\eta_p^2 = 0.016$. In term of Ne/ERN (CRN), significant force effect was observed between the Ne/ERN peak amplitude (57.16 ± 3.59 $\mu V/cm^2$) and the CRN peak amplitude (51.06 ± 3.31 $\mu V/cm^2$), $F(1,39) = 15.563$, $p < 0.001$, $\eta_p^2 = 0.291$. There was no significant force effect observed for the Ne/ERN (CRN) area, $F(1,39) = 0.637$, $p = 0.429$, $\eta_p^2 = 0.016$.

ANOVA results for the first button press showed neither significant force effect for Pe (Pc) peak amplitude, $F(1,39) = 0.459$, $p = 0.502$, $\eta_p^2 = 0.012$, nor for Pe(Pc) area, $F(1,39) = 2.537$, $p = 0.119$, $\eta_p^2 = 0.063$. As for Ne/ERN (CRN), significant force effect was observed for the first button press between Ne/ERN peak amplitude (-49.83 ± 3.44 $\mu V/cm^2$) and CRN peak amplitude (-46.83 ± 3.52 $\mu V/cm^2$), $F(1,39) = 6.313$, $p = 0.016$, $\eta_p^2 = 0.142$. Similar results pattern was observed for Ne/ERN (CRN) area, $F(1,39) = 5.994$, $p = 0.019$, $\eta_p^2 = 0.136$. The Ne/ERN area was found to be significantly higher (28.11 ± 2.03 $\mu V/cm^2$) than the CRN area (25.96 ± 2.07 $\mu V/cm^2$).

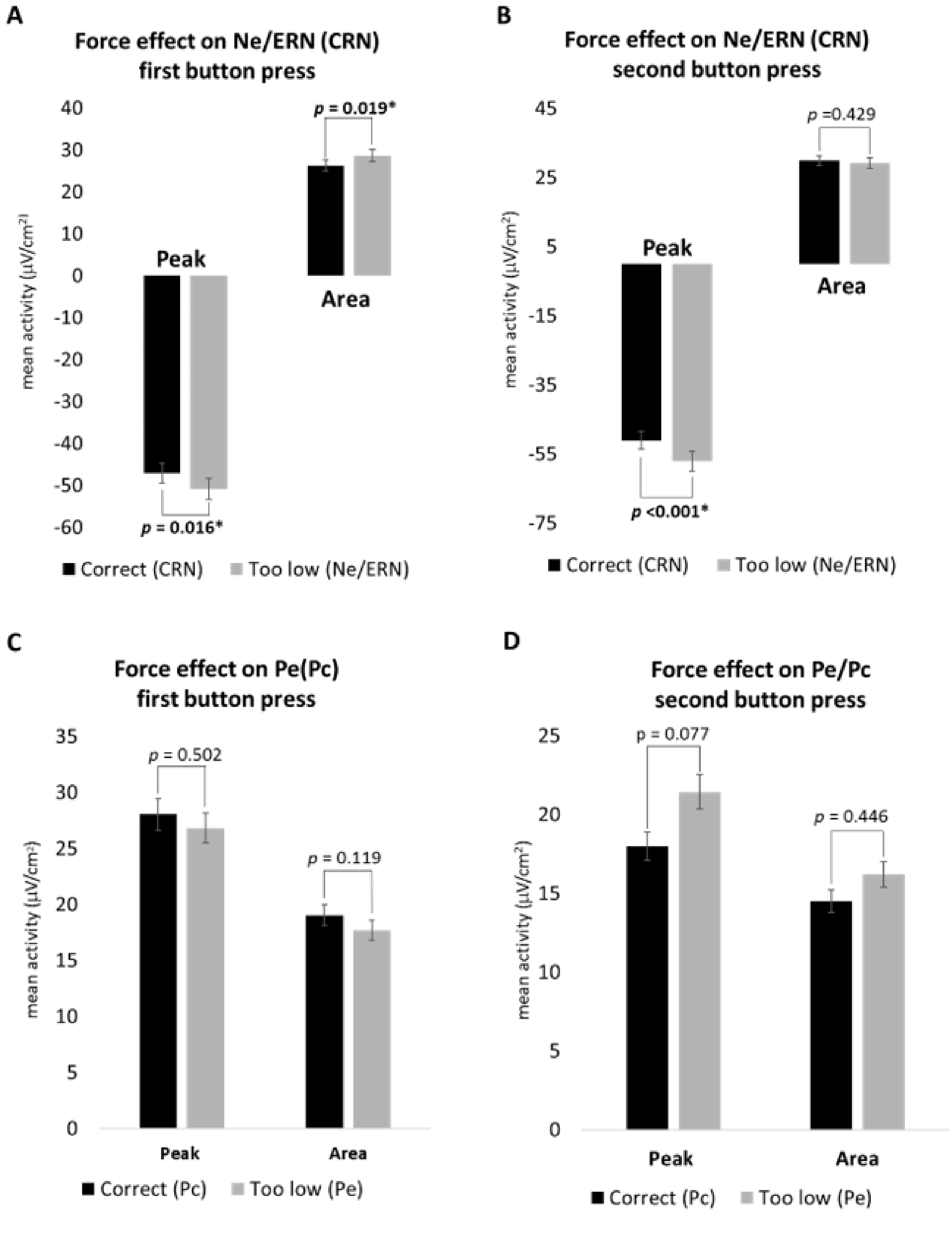


Figure 22. Force effect for the Ne/ERN(CRN) areas and peak amplitudes (A, B) and for the Pe(Pc) areas and peak amplitudes (C, D), separately shown for the first button press (A, C) and the second button press (B, D). Significant differences between the force ranges are marked with asterisks.

Force effect on Feedback-locked Averaged Event-related Potentials

A series of repeated measures ANOVAs were conducted separately on feedback-locked data for the first and second button press (i.e. left-index finger and right-index finger) to see if there are difference(s) in term of feedback processing between the two force conditions (i.e. *correct* and *too low* conditions) in term of ERP components (i.e. FRN and Feedback P3 peak amplitudes and areas) in the time window of interest, specific for each ERP component (see Figure 22, 23).

ANOVA results for the second button press exhibited significantly higher (more negative) FRN peak amplitude in the *too low* condition (-17.78 ± 2.61 $\mu V/cm^2$) than the *correct* condition (-12.14 ± 2.19 $\mu V/cm^2$), $F(1,39) = 9.080$, $p = 0.005$, $\eta_p^2 = 0.189$. Significantly higher Feedback P3 peak amplitude was also observed in the *too low* condition (33.96 ± 3.80 $\mu V/cm^2$) compared to the *correct* condition (19.03 ± 2.18 $\mu V/cm^2$), $F(1,39) = 23.270$, $p < 0.001$, $\eta_p^2 = 0.374$. Highly similar results pattern was observed on Feedback P3 area, $F(1,39) = 20.028$, $p < 0.001$, $\eta_p^2 = 0.339$.

No significant force effect was observed for the first button press between the FRN peak amplitude in the *too low* condition (-15.05 ± 2.17 $\mu V/cm^2$) and the *correct* condition (-14.47 ± 2.18 $\mu V/cm^2$), $F(1,39) = 0.239$, $p = 0.628$, $\eta_p^2 = 0.168$. However, significantly higher Feedback P3 peak amplitude for the first button press was observed in the *too low* condition (28.97 ± 3.79 $\mu V/cm^2$) than the *correct* condition (18.88 ± 2.21 $\mu V/cm^2$), $F(1,39) = 7.855$, $p = 0.008$, $\eta_p^2 = 0.168$. Highly similar results pattern was observed on Feedback P3 area, $F(1,39) = 7.038$, $p = 0.011$, $\eta_p^2 = 0.153$.

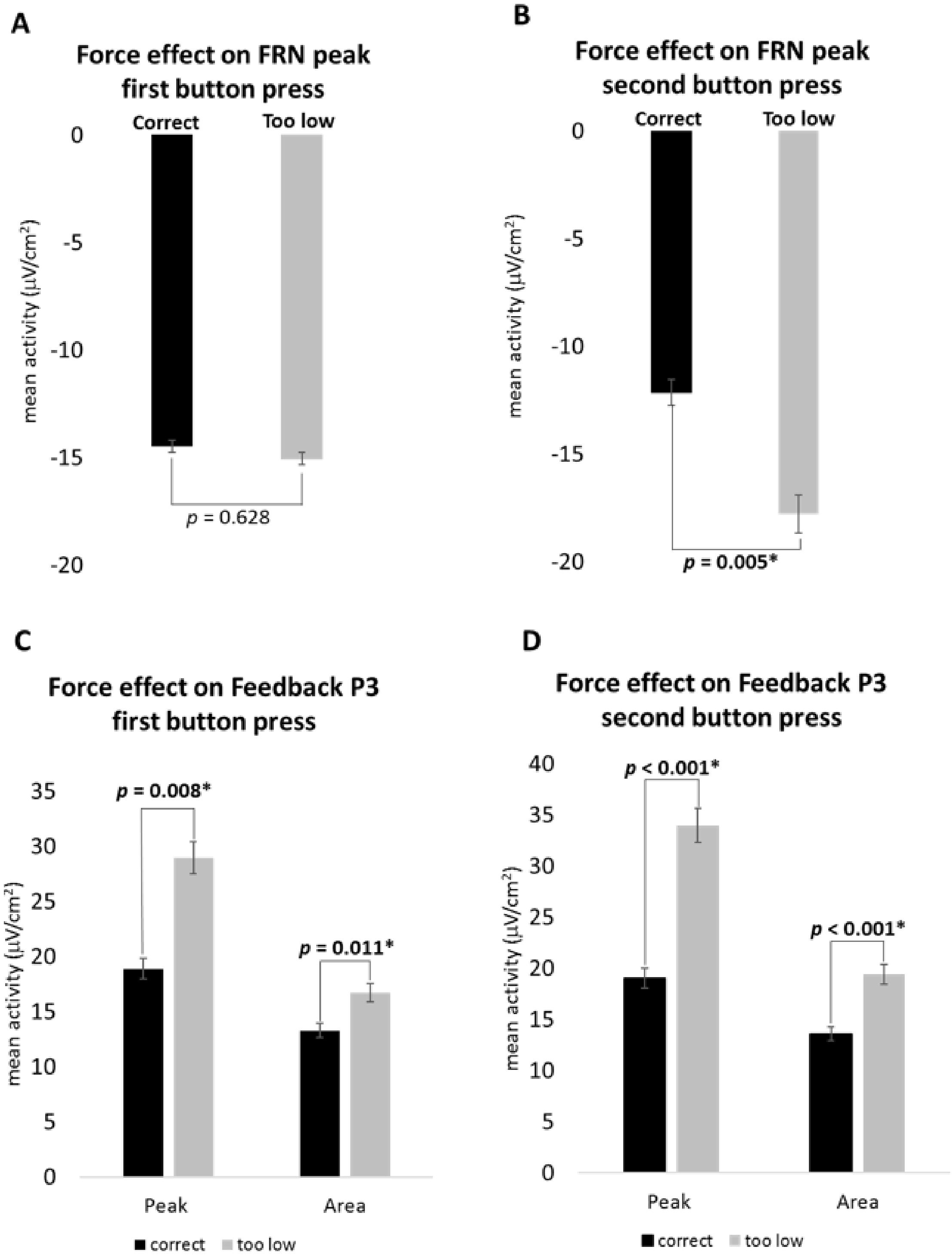


Figure 23. Force effect during feedback processing for the FRN peak amplitudes (A, B) and for the Feedback P3 areas and peak amplitudes (C, D), separately shown for the first button press (A, C) and the second button press (B, D). Significant differences between the force conditions are marked with asterisks.

2.2.4 Discussion

In the second paradigm, the PRP effect known from S1-S2 paradigms was also observed for this new R1-S2 force production paradigm. Instead of using *discrete* choice tasks, two 'continuous' force production tasks were used: the first task was done with the left-index finger and the second task with the right-index finger. In this study, not only RT but also response force parameters (i.e., PF and TTP; see Ulrich & Wings 1991), as well as the neural aspect of force monitoring (i.e., analyzed ERP components, see methods) were investigated. First, to induce a PRP effect, it is important to control the difficulty level of the two subsequent tasks, as PRP effect was reported to be diminished when there is a huge difference of difficulty between the two tasks (i.e., the first task is too easy in comparison the second task; see Ruthruff et al., 2006). Therefore, a series of pilot tests with a total of 25 participants were conducted to ensure that both tasks have a similar level of difficulty. The number of choice alternatives was also controlled to simplify the paradigm, by incorporating only one target range for both the first and second tasks, since evidence from previous literature (e.g., Jolliccoeur et al., 2002) showed that the number of response alternatives in the second task was additive with the RT of the second task. Further attempt to control for possible confound was made considering the brain asymmetry. Since the left hemisphere of the human brain controls the movements of the right body parts (and vice versa), the decision to use left-index finger for the first button press (Task 1), and the right-index finger for the second button press (Task 2) was made, to avoid bottleneck caused by unfinished motor execution, in case responses were produced by only one hand.

2.2.4.1 PRP-like Effect During Continuous Force Production: Behavioral Level

A PRP-like effect was successfully replicated on RT for the second button press. RT was the slowest in the short RSI condition, followed by the medium and the long RSI condition. This delayed RT effect is a common effect in PRP paradigms, and will be further discussed in a latter part of the discussion section (see section 2.2.4.3). Interestingly, significant RSI effect was also observed on the PF and the TTP of the second button press. The highest force was observed in the short RSI condition, followed by the medium and the long RSI condition. These results seem reasonable, since the PRP effect was most pronounced at the short RSI condition, during which the participants – who were actually not ready to produce the second response, felt the time pressure to response quickly and simultaneously reach the correct target range. The unavailability of some mental resources to monitor the second button press in a very short period of time might lead the participants to overly assign the number of force units they need to produce the force response. As for the TTP, the participants reached the peak force the fastest during the long RSI condition, because in this RSI condition they have completed the *Central Processing* for the first button press way earlier (compared to the shorter RSI conditions) and were expecting to 'complete' the second response as fast as possible. This might consequently lead to the participants producing a lower response force, since they spent less time to reach the PF (less force units were produced in a shorter time period, than in a longer time period). Further discussion regarding the PF and TTP modulations (following the delayed RT effect) will also be presented in the latter part of the discussion section (see section 2.2.4.3).

2.2.4.2 PRP-like Effect During Continuous Force Production: Neural Level

Response-locked Averages

First, in term of Ne/ERN (CRN), a highly similar negative deflection like Ne/ERN was observed in the *correct* condition (which was referred to as CRN), and a general force effect was observed between the two force conditions (i.e., *correct* and *too low*). Ne/ERN in the *too low* condition was higher than the CRN in the *correct* condition, which was a common finding in studies investigating performance monitoring: error monitoring activity is larger in the erroneous than the correct condition (e.g., Hajcak et al., 2004, Hogan et al., 2005). What is particularly remarkable, a PRP-like effect was reflected on both the Ne/ERN (in the incorrect-*too low* condition) and CRN (in the *correct* condition) of the second button press. A significantly lower Ne/ERN(CRN) in the short RSI condition indicated some kind of impaired cognitive process which resulted in lower error monitoring activity compared to the longer RSI conditions. In regard to the smaller Ne/ERN observed in the short RSI condition (compared to the other two RSI conditions), the brain seemed to lack some mental resources to perform the same amount of error monitoring activity observed in the medium and long RSI condition, presumably because the *Response Preparation* stage for Task 2 in the short RSI condition was interrupted by other processes, thus it was more difficult for the brain to retrieve an adequate amount of information (e.g., a planned amount of force units released at a certain time point) necessary for efficient error monitoring. The process of information retrieval from the *Stimulus Processing* and *Response Preparation* stages might be further interrupted because the participants were conditioned under a time pressure. What is more interesting, the absence of significant RSI effect on Ne/ERN between the medium and long RSI condition suggested that this effect was reduced (if not completely diminished) when the time interval between the two tasks was at least 700 ms (i.e., time interval for the medium RSI condition). Of course this does not exclude the possibility that there might be an 'optimal' time period between 700 ms (the medium RSI) and 1100 ms (the long RSI) in which an optimal error monitoring could be performed.

The next component that was investigated was Pe (Pc). Like in the first study, Pc (Pe in the *correct* condition) was also observed in the *correct* condition of the second study. Furthermore, an interesting RSI effect was reflected in the Pc of the second button press. Previous literature has suggested that Pc indicated participants' uncertainty on their response (e.g., Hewig et al., 2011), which could also explain the present study's results. This 'uncertainty' might be caused by an unsuccessful attempt to derive an accurate representation of the correct response (even though in this case, the correct force response might have been executed). In the present study, Pc in the medium RSI condition was significantly lower compared to the other two RSI conditions. On the other hand, Pc was highest at the short RSI condition, and was significantly higher than the medium RSI condition. This result was similar to the CRN results; higher uncertainty in the short RSI condition (as reflected by the Pc) was a product of insufficient error monitoring (indicated by the lower CRN in this RSI condition), and the Pc was significantly lower in the medium RSI condition, during which the participants presumably had enough time to retrieve as much information needed for an efficient monitoring, so that the uncertainty factor that was accounted in the short RSI condition could be avoided. However, in the long RSI condition, the uncertainty level (indicated by a higher Pc compared to the medium RSI condition) seemed to have increased again, presumably because the participants had more than enough time until they could execute their response, and this period of 'waiting' might cause some kind of interference from other cognitive processes, which consequently led to higher response uncertainty in this RSI condition.

Feedback locked averages

Separate feedback presentation for the first and second button press provided an opportunity to see how the brain processed feedback for each task. The feedback locked results (see Figure 23) showed a general force effect during feedback processing of both – the first and second button press. FRN peak amplitude (and area) appeared to be larger during the *too low* condition than the *correct* condition (for both tasks). This result is consistent with the first study's result. Like the first study, more pronounced FRN during the *too low* trials seemed to be induced by a violation of expectation. What is remarkable, this effect arose regardless the different RSI between the two tasks. To put it simply, regardless the increased level of difficulty to execute the second task caused by the different RSIs (i.e., difficulty to monitor the response quality / *correctness* for the second task was affected by the length of RSI, as exhibited in the response-locked results), the participants still hoped for a 'positive' outcome (i.e., in this case, a feedback indicating a *correct* response in every RSI conditions), which resulted in a larger FRN when this expectation is violated. This result is, like the first study, in line with the existing literature (see Holroyd et al., 2009; Hajcak et al., 2007; Holroyd & Krigolson, 2007) – in which larger FRN was observed when a response outcome was different from what was expected.

Another interesting thing to mention is, this force effect emerged not only for the second button press (which was done by the right hand – the more dominant hand for the participant) but also for the first button press (which was done by the left hand). This result seemed plausible, since the participants were instructed that both key presses are of equal importance. Note that the FRN amplitude for both force conditions appeared to be smaller for the first button press, which seemed reasonable since this response was executed by the not-dominant hand. In other words, even though the participants might expect a correct response from their left hand, this 'expectation' might not be as high as their expectation for the responses done with their more dominant hand (the right hand). Other than *handedness*, a possible reason for the generally smaller FRN for the first button press is the fact that the first button press happened much earlier in each trial, and in addition to that the feedback for the second button press was shown after the feedback for the first button press, which enhanced the likeliness that memory regarding response quality for the first button press was replaced (or *at least*, contaminated) by the second button press, which consequently led to a decrease of the general level of expectation for the first button press.

The Feedback P3 results also showed larger peak amplitude (and areas) for the *too low* condition. This effect was also shown not only during the second button press but also the first button press. This finding is in line with the existing literature, for instance, there were studies indicating an increased Feedback P3 during negative feedback as opposed to positive feedback (Groen et al., 2007; Groen et al., 2008). In the current study, a feedback showing that the participants have made a *correct* response is likely to be perceived as a 'positive' feedback, compared to a feedback indicating an error response. As a result, the more 'negative' feedback (i.e., feedback showing *too low* errors) induced higher Feedback

P3 amplitudes. Another factor that might add to the increased Feedback P3 in the *too low* condition is the frequency of occurrence of this condition. Previous studies have suggested that one of the Feedback P3's prominent features is that it decreases when target probability increases (Donchin & Coles, 1998), possibly due to an increased conformity to expectations (Hajcak et al., 2005; Yeung & Sanfey, 2004). Considering that the *too low* condition occurred less frequently than the *correct* condition, this might have caused a more pronounced Feedback P3 in the *too low* condition.

2.2.4.3 PRP effect in subsequent force production tasks: what caused the delay of Task 2's processing?

The presented paradigm demonstrated successful replication of the PRP effect on RT2, whereas RT1 was unaffected by these variables. However, there are some differences between the PRP effect observed in the current paradigm, in comparison to the PRP effect observed in the previous studies which mostly utilized two subsequent discrete choice tasks (e.g., Luck, 1998; Tombu & Jolicoeur, 2002; Ruthruff et al., 2003). The first difference is that most contemporary accounts of the PRP paradigm are primarily concerned with the nature of the central bottleneck stage during the overlap between the processing of Task 1 and Task 2, where other bottleneck scenarios of Task 2's processing received little to no attention. This is surprising, considering that some studies suggested that even when S2 is presented after R1 was executed (thus, Task 1 and Task 2 do not overlap in time, like in the current paradigm), RT2 was still found to be decreased with increasing RSI (e.g., Welford, 1980; Pashler, 1994; Jentzsch et al., 2007). To simplify, most PRP-effects observed in the literatures occurred during the condition RT1 > SOA, while in the current paradigm, the PRP effect was observed even when RT1 < RSI, possibly similar to the *residual PRP effect* observed by Jentzsch et al. (2007).

The second difference lies on the task execution stages. In discrete choice tasks, there are three stages of information processing from stimulus presentation until before the execution of motor response (as described in the literature review section): (1) *Perceptual Encoding*; (2) *Central Processing*; and (3) *Response Processing*. During the *Perceptual Encoding* stage, one tries to decode the given stimulus to retrieve necessary information for response planning. In a PRP paradigm, this stage can be manipulated to vary perceptual processing time to induce a bottleneck, for example, by manipulating the stimulus contrast (e.g., Pashler & Johnston, 1989; Jentzsch et al., 2007). Next, *the Central Processing* stage consists of reviewing the response alternatives and choosing the supposedly 'right' outcome. In this stage, the brain reviews all the available information obtained during the first stage (*Perceptual Encoding*), and decides which response is expected from the shown stimulus. The level of cognitive load during this stage would depend on the congruency of the stimulus and the expected response (i.e., cognitive load is lower when one is instructed to press the 'right arrow' on the keyboard when a stimulus is presented on right side of the computer screen, as opposed to pressing the 'left arrow' on the keyboard for the same stimulus), as well as the number of choice alternatives (i.e., it is easier and faster to decide between two response choices as opposed to four response choices). The last stage in this scheme (just before the motor execution takes place), is the *Response Processing* stage, during which the brain decides how the chosen response will be carried out (e.g., if during the *Central Processing* stage the brain decides to go for the 'right arrow' on the keyboard, in this *Response Processing* stage the brain decides which hand and which finger should be used to execute the planned response). However, when the goal of the two subsequent tasks was to produce a certain response force like in the current paradigm, the information processing stages during the task execution itself should be adapted accordingly. Since accurate force production was the main focus of the task, motor execution as well as the corresponding monitoring processes (which are necessary to ensure the force is produced as accurate as possible) should be included in the information processing stages (i.e., specifically, in the Central Processing stage). It is important to note that in the current paradigm, the participants were not presented with more than one choice alternative of response (i.e., they were required to produce force in a certain range, and this range does not change throughout the whole experiment), and the order of stimulus were always the same in every trial (i.e., left hand response followed by the right hand response).

Taking into account the differences between the information processing stages of the current paradigm and the usual PRP paradigm (in discrete choice tasks), three alternative models were contrasted. In these models, adapted versions of the regular information processing stages (specific to the current force paradigm) were proposed to illustrate the PRP effect that was observed in the current force paradigm. The information processing stages started with the *Stimulus Processing* stage, basically a less complex stage than the usual *Perceptual Encoding* stage in the PRP paradigms involving discrete choice tasks, considering that information retrieved by the participants did not vary in each trial (but rather, each stimulus presentation was just a 'response cue' for the participants). To avoid any bottleneck happening at this early stage, stimuli for the left and right hand responses were presented at the same level of contrast, and both rectangles for response cue were colored in the same color (white). To further exclude the possibility of delays induced by incongruent stimulus and response, the white rectangle appeared on the left side of the computer screen for the left hand response, and on the right side of the computer screen for the right hand response. Following the *Stimulus Processing,* the next stage is the *Central Processing for Force Task,* which is basically the 'Central Processing' stage where an overlap between tasks could not happen according to the *Central Bottleneck Theory,* since this stage introduces the highest cognitive load compared to the other stages. Note that the same term, 'Central Processing', was used for this stage for an easier comparison between the discrete and continuous task (in the context of the *Central Bottleneck Theory*). However, unlike the *Central Processing* stage in the discrete choice task, the *Central Processing* stage for the continuous force production task (as assumed by these newly postulated models) consists of two sub-stages: (1) the *Response Preparation* stage; and the (2) *Action Monitoring* stage. Unlike the discrete choice task design in which motor execution is excluded from the information processing stages, the current paradigm considers motor execution as an integrated part of the information processing stages. This consideration was made considering the fact that the 'response' itself did not end with a decision of how much Response Force should be produced, but continuous controlling, adaptation, and evaluation mechanisms were involved - at least until the Peak Force was reached, to continuously ensure that the task execution was going as planned. That being said, the *Response Preparation* (or *motor programming*, Rosenbaum 1985) stage consists initial planning regarding the force response parameters which were investigated in detail during the first study: PF and TTP. This stage is followed by the *Action Monitoring* stage, which consists of *Motor Execution* stage, during which the amount of exerted *Force Units* and the required time to reach the amount of planned response force are continuously monitored (for details, see Ulrich & Wing, 1991), and *Response Evaluation* stage, during which the brain evaluates the quality of the resulting response force. Basing on these adapted information processing stages, three alternative models aimed to explain the observed PRP effect in the current force production paradigm were postulated.

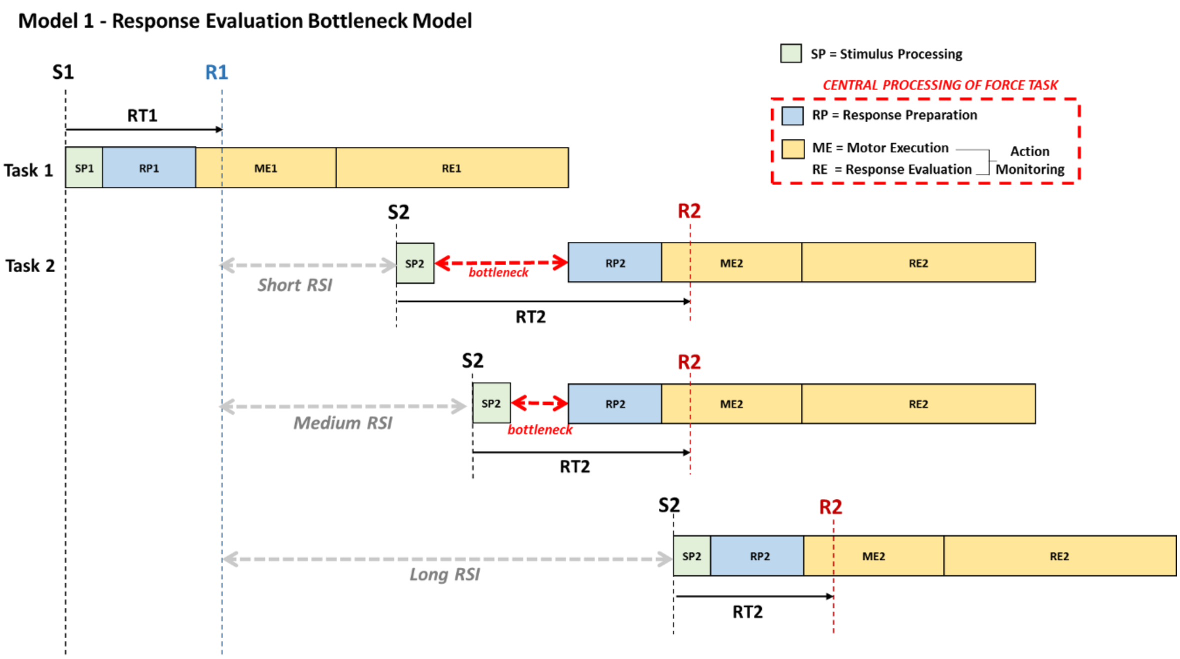


Figure 24. Response Evaluation Bottleneck Model (Model 1) for short, medium, and long RSI scenarios. The information processing stages are color-marked in: green for Stimulus Processing stage(SP), blue for Response Preparation stage (RP), yellow for Action Monitoring stage (Motor Execution and Response Evaluation). The Central Processing stage consists of Response Preparation and Action Monitoring.

Model 1: Response Evaluation Bottleneck Model

Within the first model, the *'Response Evaluation Bottleneck Model'*, it is assumed that some sort of bottleneck was induced by *Response Evaluation* stage of Task 1 (see Figure 24). This model suggested that the *Central Processing* stage for Task 2 cannot begin until the *Response Evaluation* 1 is finished, which consequently leads to a bottleneck before the *Response Preparation* 2 (as seen in the short RSI condition). In the medium RSI condition, the bottleneck still exists, but with a reduced duration, which resulted in a faster response time compared to the short RSI condition. This bottleneck effect is nullified in the longer RSI condition, during which the participants had enough time to finish the *Response Evaluation* stage of Task 1, before processing Task 2. This model is in line with a study by Welford (1980), which offered a specific explanation for factors that might contribute to a bottleneck (i.e., during a PRP paradigm when SOA > RT1). Welford proposed that an extended bottleneck period might be caused by the monitoring process for Task 1 response (R1), which requires a fresh retrieval of Task 1 Stimulus (S1) and R1 codes from the previous information processing stages. Note that in the postulated model (which was adapted to the paradigm used in the present study), the evaluation process for R1 (which requires retrieval of S1 and R1 codes) is seen as a part of the *Central Processing* stage of

Task 1, thus offering a slightly different interpretation. In Welford's model, the evaluation process of R1 (which requires retrieval of S1 and R1 codes) is seen as the bottleneck itself, while the proposed model suggested that, the *Response Evaluation 1* causes the bottleneck (instead of being the bottleneck itself), because the *Response Preparation 2* has to wait until this stage is completed. Furthermore, there could be other processes involved during this bottleneck phase, that might enhance the delay. For instance, one or more memory related processes of recalling the second response (R2) codes from the previous trial might simultaneously happen, in order to prepare for executing R2 in the corresponding trial. This interpretation is supported by evidence from the previous studies that observed *adaptation attempts* following an erroneous response (*post-error adaptations*; see Nieuwenhuis et al., 2001; Overbeek et al., 2005), in which the participants tried to incorporate the previous trial's outcome to improve response quality at the corresponding trial. Note that the application of this model might be limited since the participants were not asked to evaluate their performance for each button press. In spite of that, this model still could be used explain the behavioral data's result patterns. The bottleneck that was assumed for the short and medium RSI conditions have put the participants under time pressure, that led them to exert higher PF, compared to the long RSI condition. Further, the TTP results showed no significant differences between the medium and short RSI, but longer TTP in both conditions compared to the long RSI condition. This result might be interpreted in a way that, because of the cognitive bottleneck, the brain was somehow 'conditioned' to set a longer TTP in the short and medium RSI condition, since the participants could not assign enough force units in a very short time (see PFUM model; Ulrich & Wings, 1991), so they had to choose another strategy to exert a higher force (which is induced by the time pressure), and the only reasonable strategy to achieve this is by 'assigning' a longer TTP. On the other hand, time pressure is not a deciding factor (in term of the PF) in the long RSI condition. Therefore, the participants could control the PF better and therefore were able to reach the PF within a shorter time period. Interestingly, in term of neural variations, higher monitoring activities were observed during the medium and long RSI conditions than the short RSI condition. What is remarkable, that even though there was still a bottleneck during the medium RSI condition – which consequently caused a delayed R2 in this condition, no significant difference (in term of monitoring activities, as reflected by Ne/ERN and CRN) was observed between the medium and long RSI, which suggested that the brain was still capable of performing more or less similar monitoring activity even in a bottleneck condition, given enough time. Furthermore, the Pc results indicated less uncertainty in the medium RSI condition compared to the long RSI condition, which could offer another interesting interpretation, that there might be an optimal time interval between the medium and long RSI used in the current study, which is 'free' from the bottleneck effect and at the same time does not require the participants to wait to execute their response when they are ready to respond. This period of 'waiting' during the long RSI condition might cause the monitoring activity to be interfered by other cognitive processes (e.g., memory retrieval, decision evaluation, et cetera), which might enhance the participants' level of uncertainty during this waiting period.

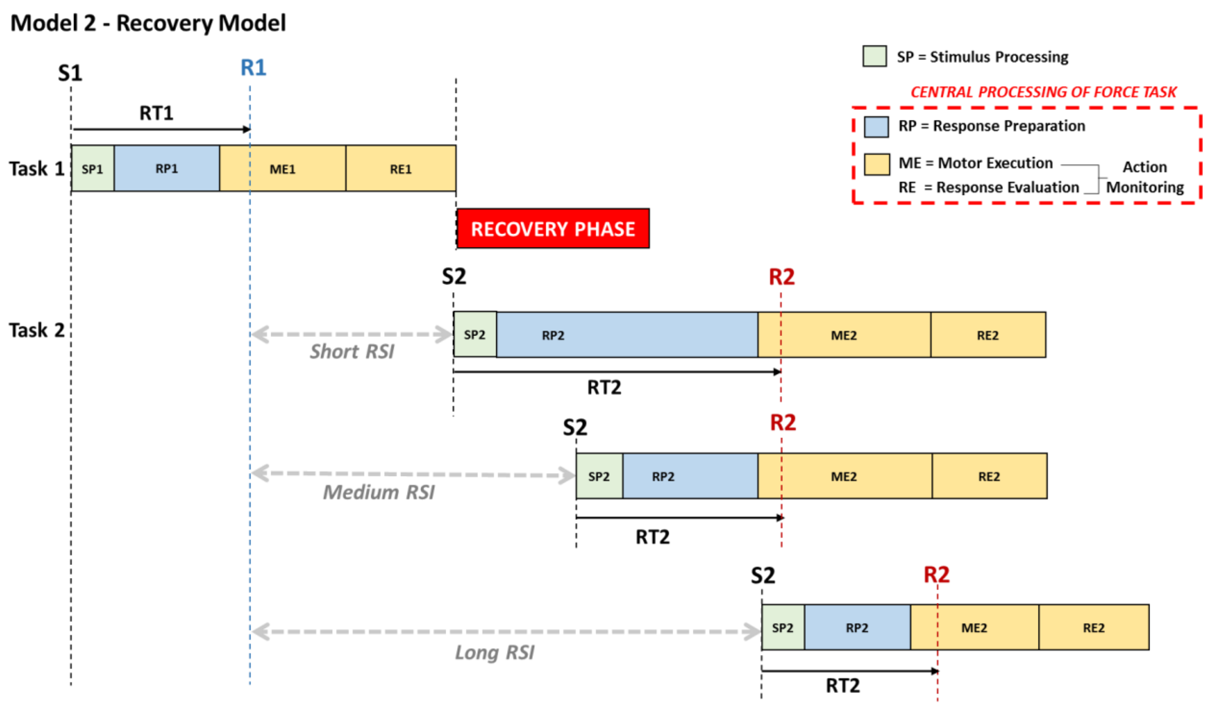


Figure 25. Recovery Model (Model 2) for short, medium, and long RSI scenarios. The information processing stages are color-marked in: green for Stimulus Processing stage (SP), blue for Response Preparation stage (RP), yellow for Action Monitoring stage (i.e., Motor Execution and Response Evaluation). The Central Processing stage consists of Response Preparation and Action Monitoring.

Model 2: Recovery Model

The second postulated model, the '*Recovery Model'*, as depicted in Figure 25, assumed that the brain needs a certain period time (i.e., a 'recovery phase') to 'reboot' before it can process another task optimally. In the short RSI condition, the *Central Processing* for Task 1 (which includes the *Response Evaluation 1*) finished before the onset of the stimulus for Task 2. However, the brain is, as assumed by this model, still at a 'recovery phase', at the same time when task 2 begins, which leads to an extended *Response Preparation* stage of task 2, caused by the lack of available cognitive resources to perform this stage at its ‘optimal’ speed. Note that the *Stimulus Processing 2* is not affected by this recovery period since this process was just a *go signal* for force execution in the current paradigm, and therefore does not inflict a considerable amount of cognitive load that might lead to an extension of this stage. This ‘recovery phase’ theory is supported by a study by Telford (1931) in which the participants had to perform a keypress response to an auditory stimulus. In this study, the stimuli were presented continuously at 0.5s, 1s, 2s, or 4s intervals according to a predefined chance order. A dramatic increased of response time (335 ms) for the 0.5 s inter-stimulus interval was found compared to the other longer intervals (241 ms, 245 ms, and 276 ms). These results pointed out the existence of a refractory period that happens after one has responded to a specific stimulus during which no other action can be executed, which is similar to the refractory period found in single neuron. This refractory period is seen a period of *intrinsic unreadiness* to emit a response which exists in a neural level, as suggested by the dramatic increase in RT presented during very short interval compared to the other longer intervals in Telford’s study. With that being said, the *Central Processing 2* stage in the long RSI condition (in the postulated model) is not affected by this recovery phase, thus, the response preparation stage can proceed without interference from the rebooting phase of all cognitive activities. Like the first model, this model could explain the behavioral results pattern found in this study. In the short RSI condition, the second stimulus appeared when the recovery phase just started, and all cognitive activities began to tone down / reboot, causing the *Response Preparation 2* stage to be extended. The participants were, at this point, aware that the second response should be produced, but unable to execute a fast response since their brain was still in the early stage of this period *of intrinsic unreadiness*. This inability to respond fast created a time pressure, which led to a higher PF in this condition. Additionally, to compensate for the lack of cognitive resources during the recovery phase, the brain deliberately set a longer time to reach the PF, as indicated by the higher TTP. On a neural level, monitoring activity for the second task was disrupted by this recovery mechanism and therefore was significantly lower than the other RSI conditions, as indicated by a low Ne/ERN and CRN. In the medium RSI condition, the brain almost completely ‘recovered’ when the *Response Preparation* stage for Task 2 started. However, this stage was still interfered by the recovery phase, thus the motor programming was done faster (compared to the short RSI condition), and more activities in term of error monitoring was observed (indicated by a higher Ne/ERN and CRN), but the participants were still under time pressure, which resulted in a higher PF and TTP compared to the long RSI condition. Meanwhile, the time pressure factor was completely eliminated during the long RSI condition, which explained the fastest RT, lowest PF, and lowest TTP for Task 2, compared to the other two RSI conditions. Interestingly, the Pc result suggested that, in the long RSI condition, the participants were almost as uncertain about their response as in the short RSI condition (i.e., Pc was lowest at the medium RSI condition, and no significant difference was observed between Pc in the short and long RSI condition). If the recovery phase was not the trigger of such uncertainty (since the recovery process ended long before the *Stimulus Processing* for Task 2 started in the long RSI condition) then a possible cause of an increased uncertainty in this condition was interference from other cognitive processes (e.g., memory retrieval of Task 1's response) which might ensue from a long period of 'waiting', like in the first model. Note that there is a conceptual difference between model 1 and model 2. Model 1 does not assume the existence of a recovery phase, thus the 'bottleneck' in model 1 is induced simply because the *Central Processing* for Task 2 cannot start until the *Central Processing* for Task 1 is completed. So, no cognitive processes were toned down (in model 1), and the delay was caused merely by a period of waiting. Unlike model 1, the *Response Preparation* for Task 2 could start when the brain is 'recovering', but since cognitive resources are limited during this recovery phase, this stage takes longer to complete, causing a delay to occur.

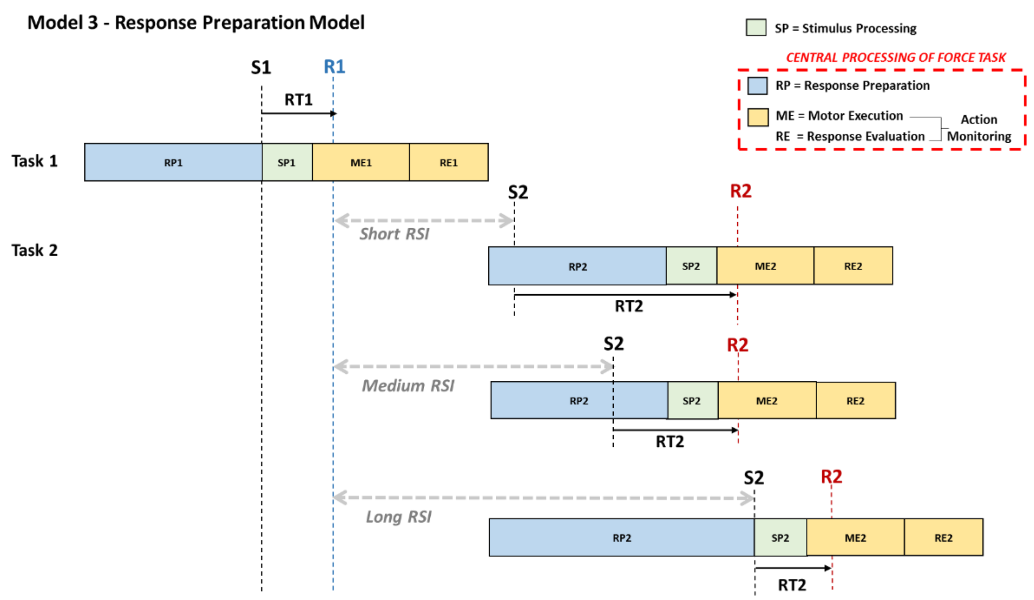


Figure 26. Response Preparation Model (Model 3) for short, medium, and long RSI scenarios. The information processing stages are color-marked in: green for Stimulus

Processing stage(SP), blue for Response Preparation stage (RP), yellow for Action Monitoring stage (i.e., Motor Execution and Response Evaluation). The Central Processing stage consists of Response Preparation and Action Monitoring.

Model 3: Response Preparation Model

The third model, the *'Response Preparation'* model (see Figure 26), offers a slightly different interpretation from the other two models, as it places more emphasize on the task execution details and the nature of the paradigm itself, instead of the 'traditional' information processing stages used in the previously conducted PRP studies. Considering that the stimulus for Task 1 (S1) and Task 2 (S2) was more like a *go signal* (thus, only a temporal information provider in this paradigm), the order of the information processing stages in this model was slightly changed. In this model, the *Response Preparation* stage is assumed to precede the *Stimulus Processing* stage, since the presented stimulus for both tasks were always the same and therefore did not affect the response decision. As illustrated in the model, the stimulus for Task 2 (S2) during the short RSI condition appeared on the screen when the *Response Preparation* for Task 2 has not yet finished, resulting in the prolongation of RT2. This prolongation of RT2 did not occur in the longer RSI condition, since the *Response Preparation* stage for Task 2 has been finished before the onset of S2. Thus, this model proposed a concept where RT2 prolongation was possible without the necessity of a bottleneck in both tasks. On a behavioral level, this model could explain the RSI effect observed in the PF (for task 2: PF during short RSI > PF during medium RSI > PF during long RSI). In the short RSI condition, the participants have not finished preparing their response by the time S2 appeared on the screen, which positioned them under a time pressure, and caused them to produce higher amount of force. Meanwhile, this time pressure situation did not occur in the long RSI conditions, in which the participants finished preparing their response and had 'spare time' to wait for the *go signal* for Task 2 (S2) to appear on the screen. Furthermore, if it was true that during the short RSI condition the participants were put into some sort of time pressure since they have not finished the *Response Preparation* for Task 2, this could also explain the Ne/ERN and CRN results pattern in this study (i.e., lower Ne/ERN and CRN in the short RSI condition compared to the other two RSI conditions, resulted from the accentuated time pressure during the short RSI condition). This explanation is in line with the evidence from the literature, in which decreased / diminished Ne/ERN during speeded tasks was reported (e.g., Dudschig & Jentzsch, 2009). Thus, following this theory, starting from the medium RSI condition, the brain presumably had enough time to produce a similar amount of monitoring activities (as reflected by the absence of significant differences of Ne/ERN and CRN between the medium and long RSI conditions). Note that in the long RSI condition, the *Response Preparation* stage for Task 2 is prolonged until the second response is cued, introducing an enhanced possibility for an 'interference' from the other cognitive activities (e.g., recalling response codes for Task 2 from previous trial, attempting to evaluate whether response for Task 1 was correct, etc.), which resulted in an increased uncertainty level in this RSI condition, as reflected by a higher Pc compared to the medium RSI condition.

Despite the different concepts underlying the proposed models, all theories presented in this study postulate that the bottleneck in the current paradigm could not have happened during the *Motor Execution* stage, since even during the short RSI condition where the period between RT1 and Stimulus onset 2 was as short as 550 ms, the participants had some time (see the grey area in Figure 27) until Stimulus 2 appeared on the screen.

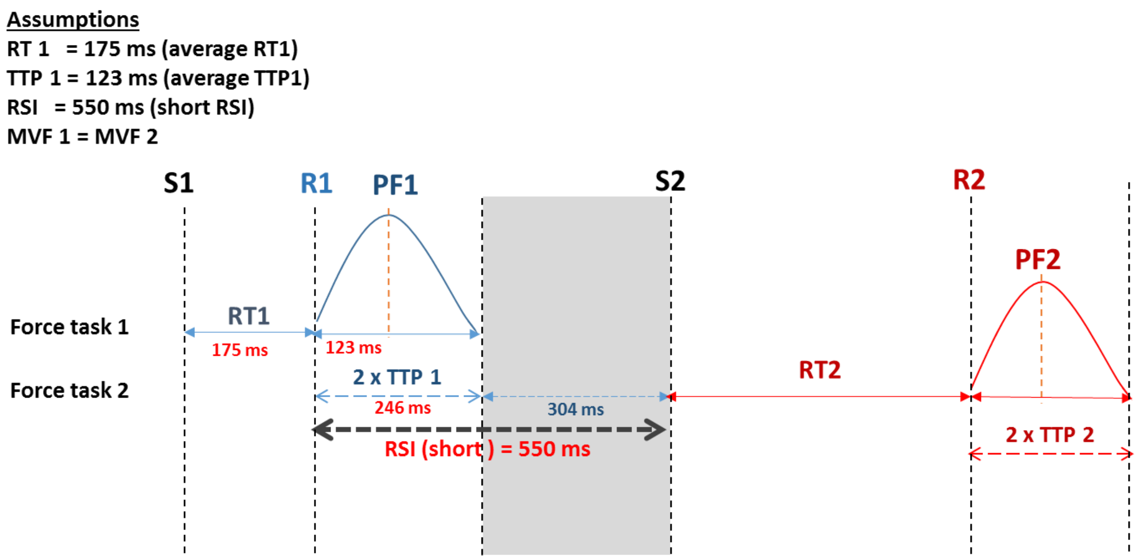


Figure 27. An example of experimental trial with a short RSI. The grey area represents the amount of time 'left' from the end of Response Execution for Task 1 until S2 onset

Thus, if a bottleneck exists, this bottleneck should have occurred after the *Motor Execution* stage of Task 1. However, one can only hypothesize what kind of processes might incur after this stage, that would cause a prolongation of RT2. It is also important to bear in mind that the postulated models rely heavily on the *Central Bottleneck Theory*, focusing on the concept that two *Central Processing* stages for two different tasks cannot happen at the same time. It is important to note that the 'Central Processing' stage in this present study was redefined (i.e., consists of *Response Execution* and *Response Evaluation*) and thus conceptually different from the 'Central Processing' stage in the discrete choice tasks (i.e., reviewing response alternatives and selecting the appropriate response). However, the original 'idea' from the *Central Bottleneck Theory* (i.e., that Central Processing stage from two different tasks cannot overlap) was taken into account while developing the models in the present study. Furthermore, it is important to consider that the participants were instructed to place equal emphasis on each of the two tasks, and to respond as fast as possible not only for the first but also the second stimulus. This instruction was intended to encourage the participants to emit their response for Stimulus 1 as soon as possible, so that the participants did not withhold their Task 1 response until their Task 2 response was also ready, consequently diminishing the possibility for a shared central capacity between the two tasks to happen – which simplified the paradigm and allowed for easier data interpretation.

Taken together, findings from the second study provided the first evidence that some kind of PRP effect is present in subsequent simple force production tasks – even when there is no motor execution overlap between the first and second task. However, further investigation would be necessary to determine the underlying cause of this effect and to disentangle the postulated models. For example, readiness-potential (RP), which was found to reflect action preparation or general anticipation of the occurrence of an action (e.g., Mele, 2017), could be investigated using stimulus-locked neural correlates to find out if the *Response Preparation* stage of Task 2 preceded the *Stimulus Processing* stage of Task 2. If RP is detected before S2 onset, and starts earlier during the long RSI condition (i.e., $RPonset_{long\ RSI} > RPonset_{medium\ RSI} > RPonset_{short\ RSI}$), this result would be in favor of model 3. Otherwise, a regression-based MVPA (SVR) could be used to regress the response-force parameters (i.e., PF & TTP) separately for the short and long RSI conditions, in order to see if the results would support one of the other two models. If the SVR result showed a long period of absence of information regarding the response-force parameters (i.e., PF and TTP) after S2 onset (and before R2 onset) during the short RSI condition, this would presumably be caused by a bottleneck that happened after the *Stimulus Processing* of Task 2, during which no evidence regarding response-force parameters should be reflected in the brain activity pattern. Additionally, if the SVR result for the long RSI condition showed an earlier prediction of the response-force parameters (in comparison to the short RSI condition), this results pattern would then support model 1. If the SVR results showed only a short period of absence of information regarding force-related parameters (which is assumed to reflect the *Stimulus Processing* stage) after the S2 onset for both the short and long RSI conditions, then model 2 should explain the PRP effect found in the present study.

2.2.5 Limitations

The paradigm used in the second study allowed for replicating PRP effect during a subsequent simple continuous tasks such as force production. Results of the data analyses have exhibited successful replication of the PRP effect on Response Time for the second force task. A similar, PRP-like effect was also observed on the response-force parameters (i.e., PF and TTP), as well as different neural correlates of force monitoring. However, the results from the *too high* condition had to be omitted since there was not enough trials for this condition. The complexity of the task itself – combined with the various attempts to control for potential confounds - made it difficult to induce an adequate amount of both *too high* trials and *too low* trials for every participant, since they were not given a specific instruction to use a certain strategy to reach the correct target force range. Suppose an equal amount of *too high* and *too low* trials could be induced for most participants, it would have been possible to see if PRP-like effect arose in all force conditions, and *error-specific differences* caused by different RSI could be thoroughly investigated. Another possible improvement to the current method is to add one or more RSI points to induce smoother curves for the PRP effects, which would allow better comparisons between the RSI conditions, that might be very useful to understand the processes involved during the information processing stages of both tasks, since these processes vary according to the RSI. However, doing this could introduce some consequences, such as: (1) it would be more difficult to get an adequate amount of trials for each force level conditions in each RSI variations; (2) this could result in prolongation of the experimental session, and thus increasing the risk of having other confounds like mental fatique.

Second, general force effects were observed during feedback presentations for both key presses. A comparison regarding PRP effect could have been conducted between the feedback for the second button press and the feedback for the first button press. However, a direct comparison (between the two feedbacks) was not possible because of an unplanned structural limitation of the current paradigm; instead of appearing always at the same time point in each trial, feedbacks for the first button press were relative to the different RSI, making it difficult to see if any effect observed during feedback presentation for the second button press was solely caused by the RSI manipulation. Suppose that an RSI effect was observed on the second button press, it would still be hard to distinguish, if this effect was induced by the different RSI or was just a general effect, since the feedback for the first button press was not completely free from the influences induced by the different RSIs.

III. General Discussion

3.1 General Perspective on a Simple One-dimensional Force Production

In the following, general aspects of force production and its monitoring based on the empirical evidence will be discussed. Firstly, I will highlight the substantial differences between force production tasks and discrete choices task. Then, findings from each study, as well as how the two studies influence each other, will be elaborated. Finally, future directions that ensue from the present studies' findings will be discussed.

As previously mentioned, it is important to fully understand the nature of the force production task used in the present study, and understand how the 'unique' nature of force production influenced the design of the paradigms used in the first and the second study. The first important concept in the force construct is the so-called *'response quality'*. As previously mentioned, response quality – in a large body of research - has been often dichotomized into *correct* and *incorrect*. The force paradigm introduced in the first and second study has not only offered a broader view on the complexity of 'simple' continuous task like one-dimensional force production, but also a different perspective to the word 'correctness'. First, this study introduced a broader definition of the term *accuracy*, which is often used as performance measure during performance tasks. The term 'accuracy' in this study is defined as an 'area' between an upper and lower limit on a force continuum. Thus, many force response outcomes can be categorized as 'correct' (and vice versa), which is a concept that is not usually applicable in most discrete performance tasks (e.g., in *n*-choices flanker tasks). The next important concept that needs to be kept in mind in the context of force production, and at the same time the most distinctive 'feature' of a force task, in comparison to the usual *discrete* task, is the *continuous* nature of this task. To produce a force response, one needs to control two aspects: (1) the time needed until the Peak Force is reached / TTP; (2) the number of force units that will be assigned continuously during this designated time period to produce a force as accurate as possible (see Ulrich & Wings, 1991 for details). This unique and continuous mechanism of force production has influenced how a number of cognitive activities were carried out (in comparison to discrete choice tasks), for example: response planning, error monitoring, and feedback decoding.

Therefore, two investigations were carried out in this study, to see how the previously mentioned cognitive processes (e.g., error monitoring) proceed during a simple force production. Note that there has not been a large body of research investigating continuous task (i.e., force production) like in the present study, therefore this section is more focused on presenting findings from the two studies in an elaborate manner, rather than comparing the present study's result with the existing literature.

The first study showed that although one cannot fully disentangle force magnitude from response quality (*correctness*), the process of determining force magnitude (i.e., *motor programming*, Rosenbaum, 1985) seems to precede the process of determining the *correctness* of a force response. Interestingly, although the brain appeared to be fast enough in term of planning the response parameters (i.e., PF and TTP) while the *correctness* aspect of the response itself was not foreshadowed in the neural signals (as shown in the first study result), it is very difficult (if not impossible), to pinpoint the exact amount of force (in cN) produced in such simple force production task. This finding also led to another interesting question: if such specific information like response force parameters (i.e., PF and TTP) was reflected in the brain activity before response onset, should not information about *correctness* be reflected somewhere in the brain activity as well? As an illustration, evaluating whether a force response is *correct* or *incorrect* (i.e., over/under-produced) should be 'easier' than stating the exact amount of produced-force (e.g., 246 cN). The latter is supposedly hard even for motor-experts. However, further investigation would be needed to investigate if information regarding *correctness* was not reflected in the brain activity (of naïve participants), or if the current method was not sensitive enough to detect evidence regarding *correctness* in the brain activity. While evidence regarding *correctness* might not be reflected in the brain activity of naïve participants because they would need more evidence to establish a mental representation of a response quality, this might not be the case with motor-experts. Unlike naïve participants, motor-experts might not require 'further evidence' (e.g., evidence provided by external feedback) to establish a mental representation of *correctness*, which consequently could be reflected in their brain activity. What is also particularly interesting to investigate, is whether different types of feedback (other than visual feedback) like sound or haptic feedback would help the brain to create a representation of *correctness* of a response. It is imaginable that, given adequate external feedback (sound, haptic, or both), a motor expert or even a naïve participant that receives enough training could precisely identify the quality of a force response.

Findings from the first study provided important evidence regarding response-dynamics in force production, and eventually led to further interesting questions, such as:

if decision regarding force response parameters was reflected as early as 270 ms before response onset (as indicated from the first study's result), would this decision process be interrupted or delayed if one had to do another force task with the same difficulty within a short period of time? Would one have enough time to evaluate the quality of the first response without the help of external feedback? Would the brain then use different strategies to program the response force parameters?

The second study served not only as an 'overview' of the course of subsequent simple force productions, but was also intended to provide insights on how the information processing stages in two subsequent continuous tasks were modulated, depending on the pause (i.e., RSI) between the two tasks. Like the *accuracy* concept in the first study, the paradigm used in the second study introduced necessary modifications (i.e., the use of RSI instead of SOA) of an existing paradigm. The results from this study have provided insights regarding response dynamics in two subsequent force production tasks. For example, in line with the first study's result that indicated evidence of early TTP programming (reflected in the brain activity pattern), the second study's results showed that no significant difference in term of TTP was observed during the short and medium RSI condition, which could be linked to the first study's finding. Since TTP was presumably programmed earlier than the PF (as indicated in the first study's result), and the *Response Preparation* stage of Task 2 (including the programming of PF and TTP) of the short and medium RSI conditions (but not the long RSI condition) was interrupted by some kind of bottleneck, it was presumably harder for the participants to make significant alteration of TTP during the motor programming of Task 2, thus varying the amount of force units might be a better 'strategy' as they were programmed later (than the TTP). It is particularly noteworthy to mention that the second study serves as the first evidence that a PRP effect exists in two continuous tasks without neither the necessity of motor execution overlap of the two tasks nor the need of increasing Task 1's difficulty level. For instance, some PRP studies in discrete choice tasks deliberately administered a difficult first task, or increased the number of response alternatives for Task 2, to induce a PRP effect. The time interval for the short RSI condition in this study could have been made shorter (i.e., to fulfill the 'usual' PRP requirement: RSI < RT1). However, with these 'arrangements', one cannot be sure if the PRP effect was induced purely because of a mental bottleneck, or if it was an 'added' effect of motor execution overlap.

One particular challenge when designing the second study's paradigm was to incorporate response evaluation for both tasks in the paradigm. In the first set of the pilot test conducted for the second study, the pilot participants were actually asked to rate their responses for both button presses, but this procedure introduced more cognitive load which exhausted the participants, rather than providing useful information. Of course, it would have been interesting to obtain information on whether the participants were aware of error commissions, which would give more information regarding feedback processing and/or response adaptation. But in exchange, incorporating response evaluation to the paradigm could introduce further confounds (i.e., mental fatigue caused by excessive cognitive load). Therefore, for the sake of simplicity and easier data interpretation, a decision to exclude this evaluation process was made.

To conclude, some important findings from both studies that could be underlined for future studies are: (1) motor programming for simple force production task was done before response onset, which might be useful for early error detection; (2) the information processing stages (i.e., including cognitive processes such as error processing, response parameters planning) in two subsequent simple force tasks are modulated by the length of RSI between the first and second task, and the differences in the information processing stages are reflected not only on a behavioral level, but also on a neural level.

3.2 Future Research

Apart from the inherent limitations in the present study, there are exciting possibilities for future research that ensue from the present findings. In regard to the first study, a possibility for future research would be, to have the participants rate their force response as *correct* or *incorrect* after each response and before feedback presentation. This mechanism would allow *detected* and *undetected errors* to be separated, whose differences might provide further insights regarding error detection related processes during simple force production. This allows to reveal whether information regarding *correctness* is available in trials with *detected* errors. Since the participants had only little time to practice before performing the task, it is interesting to see whether it is possible to detect *incorrect* force production after more practice (by using several practice sessions or by investigating motor experts such as pianist; Parlitz et al., 1998) and if this would be reflected in a clearer differentiation between the early brain activity (CRN, Ne/ERN) of *correct* and *incorrect* responses. Another possible alteration of the currently used paradigm was to vary TTP experimentally (i.e., short and long TTP; similar to Armbrecht et al., 2013). This would allow investigating how the timing of the PF might serve as an aggravating or a supporting factor for successful force error detection.

The results from the second study also provide interesting directions for future investigations. First, a multivariate approach such as MVPA could be utilized to detect what kind of information was contained in the brain activity, especially during the time interval that occurred after the first button was pressed until before the second button was pressed. This would allow determining what kind of information was processed during this period and might be a usable tool to explain the postulated models. It would be also helpful to administer questionnaires which contain different questions regarding strategies used by the participants to improve their response quality specifically for each button press. Another possibility for future research is to conduct a follow-up study on motor experts (e.g., pianists, footballers, kickboxers, etc.), which allowed more task variations that might help elucidating the underlying cause of the PRP effect in subsequent force tasks.

IV. Summary

Overall, the two studies focused on action monitoring during simple force production, a topic which has not been fully elucidated and thus needing further investigations. The first study was aimed to shed light on how force execution of a simple one-dimensional force production task unfolds. Using a modified paradigm based on the existing literature, information regarding response dynamics during a simple force execution task was gathered. Furthermore, an attempt to elucidate the time course of information relevant to the monitoring processes was made. The investigation was conducted using classical ERP components to see error-specific variations during simple force production task, and a multivariate approach was used to distinguish two different experimental conditions (i.e., different force levels and *correctness* of the response), and to predict when information regarding response force parameters such as PF and TTP are available in the brain activity pattern. The ERP results showed no clear error-specific variations between different force production ranges (characterised as *too low*, *correct*, and *too high* with respect to the target range). On the other hand, the multivariate approach (MVPA) was able to distinguish response quality (e.g., *correct* vs. *incorrect* response), as well as different force conditions (e.g., *over-produced* vs. *under-produced* force). Moreover, this approach could decode single-trial response parameters like PF and TTP before the response onset, indicating that the magnitude and timing of the force pulse were defined before response execution, while the response quality (i.e., *correctness*) of a response was not yet foreshadowed in the ERP signals.

The second study incorporated a PRP paradigm adapted to two consecutive one-dimensional force production tasks. Beyond replicating the widely known PRP-effect in discrete choice tasks, this study was aimed to see if response execution during two subsequent force production tasks was modulated by a similar PRP effect found in two subsequent discrete tasks. The investigation was carried out to see variations caused by the different length of RSI, not only on a behavioural level (i.e., by looking at RT, force response parameters such as PF and TTP, and error rates) but also on a neural level (i.e., by investigating RSI-specific variations on ERP components). On a behavioural level, the carried out investigation exhibited PRP effect not only on RT (for Task 2) but also on the response force parameters (i.e., PF and TTP) of Task 2. Moreover, PRP effect was also observed on a neural level, as reflected in the neural correlates of error monitoring (e.g., lower Ne/ERN and higher Pc in the short RSI condition). To further understand how PRP effect modulates the flow of information processing stages in two subsequent force tasks, three models based on the *Central Bottleneck Theory* were postulated. Taken together, findings from the second study provided the first evidence that a PRP effect is actually present in subsequent simple force production tasks, without the necessity of motor overlap between the first and the second task. Moreover, the different length of RSI between the two force tasks modulated the information processing stages of the second task, and the difference between RSI conditions is reflected not only on a behavioral level, but also on a neural level.

V. References

Armbrecht, A. S., Gibbons, H., & Stahl, J. (2012). Monitoring force errors: Medial-frontal negativity in a unimanual force-production task. *Psychophysiology*, *49*(1), 56-72.

Armbrecht, A. S., Gibbons, H., & Stahl, J. (2013). Effects of response force parameters on medial-frontal negativity. *PLoS One*, *8*(1), e54681.

Barbur, J. L., Watson, J. D., Frackowiak, R. S., & Zeki, S. (1993). Conscious visual perception without VI. *Brain*, *116*(6), 1293-1302.

Blackwood D. H. R, Muir W. J. (1990). Cognitive brain potentials and their application. *The British Journal of Psychiatry*, *157*, 96–101.

Bode, S., Sewell, D. K., Lilburn, S., Forte, J. D., Smith, P. L., & Stahl, J. (2012). Predicting perceptual decision biases from early brain activity. *The Journal of Neuroscience, 32*(36), 12488-12498.

Bode S, Feuerriegel D, Bennett D, Alday PM (2019). The Decision Decoding ToolBOX (DDTBOX): a multivariate pattern analysis toolbox for event-related potentials. *Neuroinformatics*, *17*(1), 27-42.

Bode, S., & Stahl, J. (2014). Predicting errors from patterns of event-related potentials preceding an overt response. *Biological Psychology, 103*, 357-369.

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, *27*(7), 1823-1835.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST),* 2(3), 27.

Cohen, M. X. (2017). Where does EEG come from and what does it mean? *Trends in Neurosciences*, 40(4), 208-218.

Cordo, P. J. (1987). Mechanisms controlling accurate changes in elbow torque in humans. *Journal of Neuroscience, 7*(2), 432-442.

Chan, A. W., Jiang, J., Chen, Y., Li, C., Prucha, M. S., Hu, Y., ... & Li, X. (2015). Progressive cognitive deficit, motor impairment and striatal pathology in a transgenic Huntington disease monkey model from infancy to adulthood. *PLoS One, 10*(5), e0122335.

Cohen, M. X. (2017). Where does EEG come from and what does it mean?. *Trends in neurosciences*, *40*(4), 208-218.

De Bruijn, E. R., Hulstijn, W., Meulenbroek, R. G., & Van Galen, G. P. (2003). Action monitoring in motor control: ERPs following selection and execution errors in a force production task. *Psychophysiology, 40*(5), 786-795.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*(5), 303-305.

Desmedt, J. E. (1982). Size principle of motoneuron recruitment and the calibration of muscle force and speed in man. *Advances in Neurology, 39,* 227-251.

Dudschig, C., & Jentzsch, I. (2009). Speeding before and slowing after errors: Is it all just strategy?. *Brain research*, *1296*, 56-62.

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology, 51*(2), 87-107.

Fahrenfort, J. J., Grubert, A., Olivers, C. N., & Eimer, M. (2017). Multivariate EEG analyses support high-resolution tracking of feature-based attentional selection. *Scientific reports, 7*(1), 1886.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495-501.

Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The error-related negativity (ERN/Ne). *Oxford handbook of event-related potential components*, 231-291.

Geisser, S., & Greenhouse, S. W. (1958). An extension of box's results on the use of the *F* distribution in multivariate analysis. *The Annals of Mathematical Statistics*, *29*(3), 885-891.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artefact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468-484.

Groen, Y., Wijers, A. A., Mulder, L. J., Minderaa, R. B., & Althaus, M. (2007). Physiological correlates of learning by performance feedback in children: a study of EEG event-related potentials and evoked heart rate. *Biological Psychology*, *76*(3), 174-187.

Groen, Y., Wijers, A. A., Mulder, L. J., Waggeveld, B., Minderaa, R. B., & Althaus, M. (2008). Error and feedback processing in children with ADHD and children with Autistic Spectrum Disorder: an EEG event-related potential study. *Clinical neurophysiology*, *119*(11), 2476-2493.

Hajcak, G., McDonald, N., & Simons, R. F. (2004). Error-related psychophysiology and negative affect. *Brain and Cognition*, *56*(2), 189-197.

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.

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.

Haynes, J. D., & Rees, G. (2006). Neuroimaging: decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*(7), 523.

Heil, M., Wahl, K., & Herbst, M. (1999). Mental rotation, memory scanning, and the central bottleneck. *Psychological Research, 62*(1), 48-61.

Hewig, J., Coles, M. G., Trippe, R. H., Hecht, H., & Miltner, W. H. (2011). Dissociation of Pe and ERN/Ne in the conscious recognition of an error. *Psychophysiology*, *48*(10), 1390-1396.

Hogan, A. M., Vargha-Khadem, F., Kirkham, F. J., & Baldeweg, T. (2005). Maturation of action monitoring from adolescence to adulthood: an ERP study. *Developmental Science*, *8*(6), 525-534.

Hogendoorn, H., & Burkitt, A. N. (2018). Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. *Neuroimage*, *171*, 55-61.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, *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.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, *14*(18), 2481-2484.

Jasper, H. H. (1958). The 10/20 international electrode system. *EEG and Clinical Neurophysiology*, *10*, 371-375.

Jentzsch, I., Leuthold, H., & Ulrich, R. (2007). Decomposing sources of response slowing in the PRP paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 610.

Jochumsen, M., Niazi, I. K., Mrachacz-Kersting, N., Farina, D., & Dremstrup, K. (2013). Detection and classification of movement-related cortical potentials associated with task force and speed. *Journal of neural engineering*, *10*(5), 056015.

Jochumsen, M., Niazi, I. K., Dremstrup, K., & Kamavuako, E. N. (2016). Detecting and classifying three different hand movement types through electroencephalography recordings for neurorehabilitation. *Medical & biological engineering & computing*, *54*(10), 1491-1501.

Jolicoeur, P., Tombu, M., Oriet, C., & Stevanovski, B. (2002). From perception to action: making the connection-a tutorial. *Common mechanisms in perception and action*, *19*, 558-586.

Leuthold, H., & Sommer, W. (1998). Postperceptual effects and P300 latency. *Psychopyhsiology, 35*(1), 34.36.

Lien, M. C., Schweickert, R., & Proctor, R. W. (2003). Task switching and response correspondence in the psychological refractory period paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(3), 692.

Luck, S. J. (1998). Sources of dual-task interference: Evidence from human electrophysiology. *Psychological Science*, *9*(3), 223-227.

Luck, S. J. (2005). An introduction to event related potentials and their neural origins. *An introduction to the event related potential technique*, *11*.

Mele, A. R. (2017). *Aspects of agency: Decisions, abilities, explanations, and free will*. Oxford University Press.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. *Psychological review, 104*(1), 3.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375-425.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38*(5), 752-760.

Overbeek, T. J., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing: On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology*, *19*(4), 319-329.

Parlitz, D., Peschel, T., & Altenmüller, E. (1998). Assessment of dynamic finger forces in pianists: effects of training and expertise. *Journal of biomechanics*, *31*(11), 1063-1067.

Pashler, H. (1992). Attentional limitations in doing two tasks at the same time. *Current Directions in Psychological Science, 1*, 44–48.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and clinical neurophysiology*, *72*(2), 184-187.

Peterson N.N, Schroeder C. E, Arezzo J. C, et al. (1995). Neural generators of early cortical somatosensory evoked potentials in the awake monkey. *Electroencephalography and Clinical Neurophysiology*, 96:248–260.

Philiastides, M. G., Ratcliff, R., & Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *Journal of Neuroscience*, *26*(35), 8965-8975.

Reason, J. (1990). *Human error.* Cambridge university press.

Rosenbaum, D. A. (1985). Motor programming: A review and scheduling theory. In *Motor behavior* (pp. 1-33). Springer, Berlin, Heidelberg.

Ruthruff, E., Pashler, H. E., & Hazeltine, E. (2003). Dual-task interference with equal task emphasis: Graded capacity sharing or central postponement?. *Perception & Psychophysics*, *65*(5), 801-816.

Ruthruff, E., Van Selst, M., Johnston, J. C., & Remington, R. (2006). How does practice reduce dual-task interference: Integration, automatization, or just stage-shortening?. *Psychological research*, *70*(2), 125-142.

Sailer, U., Fischmeister, F. P. S., & Bauer, H. (2010). Effects of learning on feedback-related brain potentials in a decision-making task. *Brain research*, *1342*, 85-93.

Sigman, M., & Dehaene, S. (2006). Dynamics of the central bottleneck: Dual-task and task uncertainty. *PLoS biology*, *4*(7), e220.

Stahl, J. (2010). Error detection and the use of internal and external error indicators: An investigation of the first-indicator hypothesis. *International Journal of Psychophysiology*, *7*7(1), 43-52.

Steinhauser, M., & Yeung, N. (2010). Decision processes in human performance monitoring. *Journal of Neuroscience,30*(46), 15643-15653.

Steinhauser, M., & Yeung, N. (2012). Error awareness as evidence accumulation: effects of speed-accuracy trade-off on error signalling. *Frontiers in Human Neuroscience*, *6*, 240.

Teplan, M. (2002). Fundamentals of EEG Measurement. *In: Measurement Science Review*, 2, 1-11.

Tombu, M., & Jolicœur, P. (2002). All-or-none bottleneck versus capacity sharing accounts of the psychological refractory period phenomenon. *Psychological research*, *66*(4), 274-286.

Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual review of psychology*, *63*, 483-509.

Ulrich, R., & Wing, A. M. (1991). A recruitment theory of force-time relations in the production of brief force pulses: The parallel force unit model. *Psychological Review*, *98*(2), 268.

Ulrich, R., Wing, A. M., & Rinkenauer, G. (1995). Amplitude and duration scaling of brief isometric force pulses. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1457.

Ulrich, R., & Miller, J. (2008). Response grouping in the psychological refractory period (PRP) paradigm: Models and contamination effects. *Cognitive Psychology*, *57*(2), 75-121.

Ullsperger, M., & Von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage*, *14*(6), 1387-1401.

van Vugt, M. K., Simen, P., Nystrom, L. E., Holmes, P., & Cohen, J. D. (2012). EEG oscillations reveal neural correlates of evidence accumulation. *Frontiers in neuroscience*, *6*, 106.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity'specific to errors?. *Biological psychology*, *51*(2-3), 109-128.

Welford, A. T. (1952). The 'psychological refractory period'and the timing of high-speed performance—a review and a theory. *British Journal of Psychology. General Section*, *43*(1), 2-19.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*(28), 6258-6264.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological review*, *111*(4), 931.

Appendix

Appendix A

Statistical Analyses Results of ERP Components in Study 1

Table A1. Peak amplitude and areas for all ERP components (correct vs. incorrect-too high vs. incorrect- too low) in study 1

RESPONSE LOCKED				
PEAK AMPLITUDE (µV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Ne/ERN (FCz)	Incorrect – too high	-25.41± 3.26	3.08	0.050
	Correct	-21.81± 2.67		
	Incorrect – too low	-22.89 ±2.71		
Pe (Cz)	Incorrect – too high	9.39 ± 2.95	3.77	0.030*
	Correct	5.09 ± 2.19		
	Incorrect – too low	4.37 ± 1.47		
AREA (mV/cm2)				
ERP component (channel)	**area (mean ± SEM)**		**F value**	**P value**
Ne/ERN (FCz)	Incorrect – too high	16.44 ± 2.25	.86	.424
	Correct	15.82 ± 1.97		
	Incorrect – too low	15.26 ± 1.92		
Pe (Cz)	Incorrect – too high	15.42 ± 2.26	6.98	0.002*
	Correct	11.29 ± 1.40		
	Incorrect – too low	9.95 ± 1.26		
TIME-TO-PEAK LOCKED				
PEAK AMPLITUDE (µV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Ne/ERN (FCz)	Incorrect – too high	-22.96± 1.69	.59	.554
	Correct	-21.60± 2.57		
	Incorrect – too low	-22.65 ±2.57		
Pe (Cz)	Incorrect – too high	15.94 ± 2.95	7.52	<0.001*
	Correct	8.41 ± 1.93		
	Incorrect – too low	8.89 ± 1.96		
AREA (µV/cm2)				
ERP component (channel)	**area (mean ± SEM)**		**F value**	**P value**
Ne/ERN (FCz)	Incorrect – too high	12.05 ± 1.98	1.68	.197
	Correct	13.66 ± 1.77		
	Incorrect – too low	12.05 ± 1.72		
Pe (Cz)	Incorrect – too high	16.72 ± 1.85	15.56	<0.001*
	Correct	10.36 ± 1.27		
	Incorrect – too low	9.87 ± 1.24		
FEEDBACK LOCKED				
PEAK AMPLITUDE (µV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
FRN (Cz)	Incorrect – too high	-3.05 ± 1.87	6.43	0.002*
	Correct	0.45 ± 1.88		
	Incorrect – too low	-5.08 ± 1.95		

FRN (FCz)	Incorrect – too high	-4.15 ± 2.03	5.54	0.007*
	Correct	-0.31 ± 2.00		
	Incorrect – too low	-5.24 ± 1.88		

Table A2. Peak amplitude and areas for all ERP components (correct vs. incorrect) in study 1

RESPONSE LOCKED				
PEAK AMPLITUDE (μV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Ne/ERN (FCz)	Correct	-21.81 ± 2.67	0.16	0.695
	Incorrect	-22.93 ± 1.99		
Pe (Cz)	Correct	5.09 ± 2.19	1.98	0.166
	Incorrect	6.51 ± 1.49		
AREA (μV/cm2)				
ERP component (channel)	**area (mean ± SEM)**		**F value**	**P value**
Ne/ERN (FCz)	Correct	15.82 ± 1.97	0.64	0.428
	Incorrect	15.56 ± 1.37		
Pe (Cz)	Correct	11.29 ± 1.40	0.63	0.433
	Incorrect	10.65 ± 1.13		
TIME-TO-PEAK LOCKED				
PEAK AMPLITUDE (μV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Ne/ERN (FCz)	Correct	-21.60 ± 2.57	0.00	0.991
	Incorrect	-21.59 ± 1.91		
Pe (Cz)	Correct	8.41 ± 1.93	2.26	0.140
	Incorrect	9.78 ± 1.53		
AREA (μV/cm2)				
ERP component (channel)	**area (mean ± SEM)**		**F value**	**P value**
Ne/ERN (FCz)	Correct	10.66 ± 1.77	2.67	0.060
	Incorrect	11.76 ± 1.01		
Pe (Cz)	Correct	10.36 ± 1.27	0.17	0.679
	Incorrect	11.20 ± 1.06		
FEEDBACK LOCKED				
PEAK AMPLITUDE (μV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
FRN (Cz)	Correct	0.45 ± 1.88	9.01	0.004*
	Incorrect	-4.37 ± 1.44		
FRN (FCz)	Correct	-0.31 ± 2.00	5.74	0.021*
	Incorrect	-4.87 ± 1.44		

Appendix B

Statistical Analyses Results of ERP Components in Study 2

Table B1. RSI effect on peak amplitude and areas for all ERP components (separately investigated for *correct* and *incorrect-too low* conditions) for Task 2 in study 2

RESPONSE LOCKED (Task 2)				
PEAK AMPLITUDE (μV/cm2)				
Ne/ERN or CRN (at FCz)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Correct (CRN)	Short RSI	-46.26 ± 2.99	5.21	0.010*
	Medium RSI	-54.26 ± 3.69		
	Long RSI	-52.79 ± 3.87		
Incorrect – Too Low (Ne/ERN)	Short RSI	-56.26 ± 4.08	0.88	0.423
	Medium RSI	-58.97 ± 3.91		
	Long RSI	-56.24 ± 3.70		
Pe or Pc (at Cz)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Correct (Pc)	Short RSI	20.52 ± 2.65	7.11	0.023*
	Medium RSI	13.86 ± 2.38		
	Long RSI	19.50 ± 3.06		
Incorrect – Too Low (Pe)	Short RSI	25.18 ± 3.12	2.697	0.081
	Medium RSI	18.63 ± 2.53		
	Long RSI	21.05 ± 3.24		
AREA (μV/cm2)				
Ne/ERN or CRN (at FCz)	**area (mean ± SEM)**		**F value**	**P value**
Correct (CRN)	Short RSI	26.39 ± 1.69	4.94	0.012*
	Medium RSI	31.65 ± 2.44		
	Long RSI	31.42 ± 2.66		
Incorrect – Too Low (Ne/ERN)	Short RSI	26.29 ± 2.03	4.39	0.020*
	Medium RSI	31.71 ± 2.40		
	Long RSI	29.44 ± 2.19		
Pe or Pc (at Cz)	**area (mean ± SEM)**		**F value**	**P value**
Correct (Pc)	Short RSI	12.46 ± 1.10	4.37	0.023*
	Medium RSI	12.46 ± 1.10		
	Long RSI	19.50 ± 3.06		
Incorrect – Too Low (Pe)	Short RSI	17.69 ± 1.87	1.145	0.329
	Medium RSI	14.22 ± 1.98		
	Long RSI	15.57 ± 1.82		

Table B2. RSI effect on peak amplitude and areas for all ERP components (separately investigated for *correct* and *incorrect-too low* conditions) for Task 1 in study 2

RESPONSE LOCKED (Task 1)				
PEAK AMPLITUDE (µV/cm2)				
Ne/ERN or CRN (at FCz)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Correct (CRN)	Short RSI	-45.89 ± 3.36	0.89	0.420
	Medium RSI	-47.05 ± 3.51		
	Long RSI	-48.49 ± 3.91		
Incorrect – Too Low (Ne/ERN)	Short RSI	-50.64 ± 3.73	1.19	0.313
	Medium RSI	-50.51 ± 3.74		
	Long RSI	-48. 47 ± 3.28		
Pe or Pc (at Cz)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Correct (Pc)	Short RSI	28.43 ± 2.59	1.16	0.324
	Medium RSI	28.96 ± 2.87		
	Long RSI	26.69 ± 2.51		
Incorrect – Too Low (Pe)	Short RSI	26.11 ± 2.41	1.05	0.359
	Medium RSI	27.76 ± 2.93		
	Long RSI	28.69 ± 2.60		
AREA (µV/cm2)				
Ne/ERN or CRN (at FCz)	**area (mean ± SEM)**		**F value**	**P value**
Correct (CRN)	Short RSI	25.18 ± 2.04	1.63	0.210
	Medium RSI	26.21 ± 2.13		
	Long RSI	27.41 ± 2.28		
Incorrect – Too Low (Ne/ERN)	Short RSI	28.27 ± 2.20	0.45	0.956
	Medium RSI	28.16 ± 2.22		
	Long RSI	27.89 ± 2.02		
Pe or Pc (at Cz)	**area (mean ± SEM)**		**F value**	**P value**
Correct (Pc)	Short RSI	18.99 ± 1.72	1.79	0.179
	Medium RSI	20.21 ± 2.07		
	Long RSI	17.92 ± 1.82		
Incorrect – Too Low (Pe)	Short RSI	16.28 ± 1.76	1.22	0.307
	Medium RSI	17.45 ± 1.96		
	Long RSI	18.17 ± 1.91		

Table B3. Force effect on peak amplitude and areas of response-locked averages for Task 2 in study 2

RESPONSE LOCKED (Task 2 – Force effect)				
PEAK AMPLITUDE (µV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Ne/ERN or CRN (FCz)	Correct	-51.01 ± 3.31	15.56	<0.001*
	Incorrect – too low	-57.16 ± 3.59		
Pe or Pc (Cz)	Correct	18.51 ± 2.30	3.29	0.077
	Incorrect – too low	21.62 ± 2.36		
AREA (µV/cm2)				
ERP component (channel)	**area (mean ± SEM)**		**F value**	**P value**
Ne/ERN or CRN (FCz)	Correct	29.82 ± 2.05	0.64	0.429
	Incorrect – too low	29.15 ± 1.94		
Pe or Pc (Cz)	Correct	14.65 ± 1.30	0.59	0.446
	Incorrect – too low	15.59 ± 1.44		

Table B4. Force effect on peak amplitude and areas of response-locked averages for Task 1 in study 2

RESPONSE LOCKED (Task 1 – Force effect)				
PEAK AMPLITUDE (µV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
Ne/ERN or CRN (FCz)	Correct	-46.83 ± 3.52	6.31	0.016*
	Incorrect – too low	-49.91 ± 3.44		
Pe or Pc (Cz)	Correct	28.47 ± 2.56	0.46	0.502
	Incorrect – too low	27.52 ± 2.36		
AREA (µV/cm2)				
ERP component (channel)	**area (mean ± SEM)**		**F value**	**P value**
Ne/ERN or CRN (FCz)	Correct	25.95 ± 2.07	5.99	0.019*
	Incorrect – too low	28.11 ± 2.04		
Pe or Pc (Cz)	Correct	18.92 ± 1.81	2.54	0.119
	Incorrect – too low	17.30 ± 1.69		

Table B5. Force effect on peak amplitude and areas of feedback-locked averages for Task 2 in study 2

FEEDBACK LOCKED (Task 2 – Force effect)				
PEAK AMPLITUDE (μV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
FRN (FCz)	Correct Incorrect – too low	-12.14 ± 2.67 -17.78 ± 2.19	9.08	0.005*
Feedback P3 (FCz)	Correct Incorrect – too low	19.03 ± 2.18 33.96 ± 3.80	23.27	<.001 *
AREA (μV/cm2)				
ERP component (channel)	**area (mean ± SEM)**		**F value**	**P value**
Feedback P3 (FCz)	Correct Incorrect – too low	13.58 ± 1.18 19.38 ± 1.49	20.03	<.001 *

Table B6. Force effect on peak amplitude and areas of feedback-locked averages for Task 1 in study 2

FEEDBACK LOCKED (Task 1 – Force effect)				
PEAK AMPLITUDE (μV/cm2)				
ERP component (channel)	**peak amplitude (mean ± SEM)**		**F value**	**P value**
FRN (FCz)	Correct Incorrect – too low	-14.47 ± 2.18 -15.05 ± 2.19	0.24	0.628
Feedback P3 (FCz)	Correct Incorrect – too low	18.88 ± 2.21 28.97 ± 3.78	7.85	0.008*
AREA (μV/cm2)				
ERP component (channel)	**area (mean ± SEM)**		**F value**	**P value**
Feedback P3 (FCz)	Correct Incorrect – too low	13.26 ± 0.89 16.74 ± 1.49	7.04	0.011*