Cognition in Action: Error Awareness in 7 Actions-per-Second Performance

C Kalfaoğlu

University of Sheffield

Dissertation submitted for the award of PhD

2012

Cognition in Action: Error Awareness in 7 Actions-per-Second Performance

Çığır Kalfaoğlu

Dissertation submitted for the award of PhD

Department of Psychology

15 August 2012

Contents

List of Figures

List of Tables

Acknowledgements

I am grateful to my supervisors Tom Stafford and Elizabeth Milne who supported me in every step of the 3 years I have worked on this project and turned my Ph.D. into a surprisingly pleasant experience. I would also like to thank Tom Walton, the discussions with whom always challenged me to think in different ways in interpreting my results. I would like to thank all of the touch-typist participants who willingly endured my long EEG experiment and provided invaluable feedback and insights about their typing experience. Finally, I would like thank Sara Elisa Fischer, for volunteering to take part in all of my testing sessions, pilot or not, as a talented touch-typists and a dedicated participant.

Abstract

In two experiments, we examined the behavioural and electro-physiological effects of errors in touch-typing. The effect of errors on skilled actions is an under-studied area in cognitive psychology. The available evidence suggests that errors have different effects on discrete vs. skilled and continuous actions. Our primary aim was to study the behavioural and electro-physiological effects of errors, and explore any interactions between them such as event-related potentials (ERP) and error correction via the backspace. We asked touch-typists to type 100 sentences in the absence of visual feedback. We recorded electro-encephalogram (EEG) as well as typing performance of touch-typists. We analysed the data using independent component analysis (ICA), with an emphasis on the difference between correct and error key-presses as well as corrected and uncorrected error key-presses. We found that the error (corrected and uncorrected) key-presses in typing were slowed, and were followed by slowed key-presses. In the EEG record, we found a considerable increase in the power of theta oscillations (3-8Hz) as well as classic ERP findings (i.e. Error related negativity (ERN) and positivity (Pe). Importantly, these effects were much stronger during corrected errors compared to uncorrected errors. Our results suggest that even in a skilled action which involves more than 7 key-presses every second, it is possible to detect one's errors before the error action is completed, and that error correction can be predicted by the strength of error induced changes in the EEG record.

Chapter 1

Literature Review and Introduction to Performance Monitoring

Performance monitoring is a cognitive ability which is crucial for keeping our ongoing actions in line with our long term intentions. Failure to make necessary changes in our ongoing behaviour in response to our mistakes (e.g. picking up salt instead of sugar) or to the environment (e.g. using dirty mug), might lead to unpleasant consequences, for example, in simple tasks such as making tea. The consequences could be grimmer when one considers more complex tasks such as driving. Potential outcomes of not realizing you are speeding up range from loss of license, to loss of life.

Studies of performance monitoring have typically focussed on errors of performance. The tasks used to generate errors have often been those that involve simple response button actions, such as the flankers (Eriksen & Eriksen, 1974) and Stroop (Stroop, 1935) tasks. In the sections that follow, classic findings (e.g. Rabbitt, 1966a, 1966b) from these tasks, and also more recent findings from studies using more ecological tasks involving piano playing (Herrojo-Ruiz, Jabusch, & Altenmuller, 2009) and typing (G. Logan & Crump, 2011) are summarized.

The chapter is structured such that the behavioural and electro-physiological indices of performance monitoring are separated into different sections. Within each section, we describe specific parameters used to quantify performance monitoring (e.g. post-error slowing, error related negativity, etc.), and finish with a final section identifying the gaps in the literature.

Skilled vs. non-skilled actions

An important concept to consider when studying skilled actions like typing is cognitive expertise. One of the most important factors involved in gaining skill or expertise in a field is deliberate practice (Ericsson & Krampe, 1993). Skilled performers' (e.g. sports people, surgeons, typists, etc.) superiority is domain specific and doesn't usually generalize to other cognitive skills (Ericsson & Lehmann, 1996). These domain specific skills that the experts gather through deliberate practice changes the way domain-related information is coded and represented (Williams & Ericsson, 2005). Typing is no exception to this pattern. According to Gentner (1984), expertise in typing is associated with qualitatively different use of mental resources, with cognitive processes overlapping in time and a decreased load on conscious and cognitive resources. As the typing skill is acquired there is a general shift from cognitive to motor limits on performance (Gentner, 1984).

1.1 Behavioural Literature

1.1.1 Findings from Discrete Actions

An overview of the findings from behavioural literature from discrete trial tasks is summarized below. These studies show that pre-error and error responses tend to be faster than correct responses, and post-error responses are slowed down compared to post-correct responses. This literature suggests that post-error slowing is a potential marker of error effects on performance, whereas pre-error and error speeding are not necessarily caused by, but indicative of the upcoming error response, as predicted by speed accuracy trade-off. We summarize the findings that lead us to assume this interpretation below.

Error Responses

Human Error When one considers the number of possible ways in which an action can go wrong, it appears that errors can take an infinite number of forms. Reason (1990) gives the example of boiling an egg. The stages and the number of ways it can be 'bungled' are numerous. Even when one considers such simple actions as those used in discrete trial tasks (as opposed to boiling an egg), there are multiple ways in which things can go wrong (errors in perception of the stimulus, execution of the key-press in time and strength, etc.). In typing, where there are 26 letters and several frequent punctuation symbols, the number of ways one can make a mistake is much larger. This is because only one correct key can be pressed at a given moment in time, and any of the 25 letters or the punctuation mark would constitute an error. It follows that there are many instances and many different ways in which an error can be generated. However, we find that errors are not usually as frequent and their form is not as variable (Reason, 2000). For example, Salthouse (1986) suggests in his review of typing errors that a vast majority of errors take one of 4 forms (i.e. substitutions, omissions, insertions and transpositions).

The mental processes that underlying these different kinds of errors are likely to be different. Even though this is not a question we specifically pursued, it is nevertheless a relevant and interesting one. Certain errors are caused simply by the finger aiming for the wrong letter on the keyboard. These misaiming errors would be an example where wrong movement of the finger is selected. One might argue that a finger moved too far to the left and missed the target letter in doing so. For example, one of the most frequent errors that typists make is pressing the key directly adjacent to the target key (MacNeilage, 1964). However, most of these adjacent errors pressed with the correct finger: Pressing 's' is typically executed by the left ring finger, and pressing 'a' is executed by the left little finger. Such errors are called substitutions, and are examples where the wrong *letter* is selected, while the *finger selected* is right (Rosenbaum, 1991).

These errors all lead to the pressing of a key that is not necessarily present in the word to be typed. However, another interesting error happens not when a wrong letter or finger is selected, but when the temporal order of the letters is broken. Examples of this include transposition errors like ‘preson’ (instead of ‘person’). Another common error type similar to transposition errors is doubling errors (e.g. ‘bokk’ instead of ‘book’) where the ‘doubling’ action starts too early or late in the sequence, and peri-errors (e.g. ‘thses’ instead of ‘these’) (F. A. Logan, 1999). The common feature of transposition, doubling and peri-errors is that the correct letters (or the doubling component) in a word are executed, but in the wrong order. This is different to substitution errors where a letter which is not found in the word to be typed replaces a correct letter. F. A. Logan (1999) refers to the difference between these two groups of errors as substitution vs. temporal errors.

It is possible that the mechanisms that fail (lead to the error commission) in substitution errors are different than temporal errors. One possibility is that a correctly typed letter triggers the pressing of a incorrect letter which is very frequently associated with that correct letter. This would be an example of a substitution error. Another kind of error would be caused a correct pre-programmed finger movement shifting in temporal order such that it is replaced by another correct and pre-programmed finger movement. This would be an example of transposition error, where all finger movements are correctly programmed, but their timing is wrong. It is possible that selection of the wrong action (e.g. substitutions) as opposed to incorrect temporal ordering of selected actions (e.g. transpositions) are different at some level, and this may have implications for the way each type of error is detected.

Unfortunately, we were not able to distinguish between different kinds of errors with confidence due to a number of reasons. These include the fact that most of the errors are corrected immediately after the pressing of the wrong letter. This makes it very difficult to judge whether an error is caused by an omission (missing of a correct letter) or a transposition. An example is the word ‘from’. If the typist types ‘fo’ and immediately presses the backspace, it becomes difficult to judge post-hoc whether the error was a transposition between ‘o’ and ‘r’, or an omission of the letter ‘r’ without observing the rest of the letters typed.

General Error Effects on Performance To our knowledge, the earliest reports of error effects on performance were those conducted by Rabbitt in the 1960's (Rabbitt, 1966a, 1966b, 1967, 1968). These classic experiments showed that the incorrect responses were associated with faster than usual reaction times. In two experiments investigating reaction times to visual stimuli, Rabbitt showed that i) participants could correct their mistakes very fast and without external feedback and ii) reaction times in error trials were faster than those in correct trials (Rabbitt, 1966a, 1966b). The finding that participants can correct their responses very fast without external feedback (Rabbitt, 1967) has important implications for our methodology as well as our results concerning performance monitoring.

The observation of fast error responses is typically explained within the speed accuracy trade-off framework (Wickelgren, 1977) such that the faster the response, the less likely it is to be correct. This observation of error speeding has been replicated many times in studies of performance monitoring using different sorts of discrete trial tasks, such as Eriksen's Flankers task (Armbrecht, Wohrmann, Gibbons, & Stahl, 2010; H. Eichele, Juvodden, Ullsperger, & Eichele, 2010; Hajcak, Nieuwenhuis, Ridderinkhof, & Simons, 2005; Heldman, Russeler, & Munte, 2008; Maier, Steinhauser, & Hubner, 2010; Nieuwenhuis et al., 2002; Pailing, Segalowitz, Dywan, & Davies, 2002; Rodriguez-Fornells, Kurzbuch, Munte, & Munte, 2002), go/no-go task (Scheffers, Coles, Bernstein, Gehring, & Donchin, 1996) and Stroop task (Gehring & Fencsik, 2001; Hajcak & Simons, 2008) with stimuli presented in different sensory modalities (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991) and when responses required the use of different parts of the body (Gehring & Fencsik, 2001; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001; Van't Ent & Apkarian, 1999).

Even though error speeding is a robust finding in the literature, it is worth noting that it is sensitive to the methodology used. For example, one exception to this pattern was reported by de Bruijn, Hulstijn, Meulenbroek, and van Galen (2003) in a force production task. When the participants' grip on the dynamo-meter was too strong and not strong enough, their response times were slower and no different than when their force was accurate, respectively.

Partial Errors There are many studies which used more fine-grained methods including electromyography (EMG) and/or dynamo-meters to assess the temporal dynamics of error commission (e.g. Burle, Roger, Allain, Vidal, & Hasbroucq, 2008; Coles, Scheffers, & Fournier, 1995; Fournier, Scheffers, Coles, Adamson, & Abad, 1997; Gehring & Fencsik, 2001; Gehring, Goss, Coles, Meyer, & Donchin, 1993; Gehring, Himle, & Nisenson, 2000; Gehring & Knight, 2000; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000). The use of EMG allows a higher level of precision in tracking the exact time the muscles of the fingers are activated in response to a stimulus. These studies show that a number of correct key-presses are preceded by weak but reliable activations of the error response. These activations are observable on the dynamo-meter record and detectable by the EMG, but they are not strong enough to result in an overt button press (hence the name *partial* errors). Results from these studies suggest that the correct responses preceded by partial errors are slower than those which are not. Further, partial errors are more probable when the imperative stimulus contains information associated with both responses (i.e. incompatible conditions), compared to when it contains information that favours only one response (i.e. compatible conditions), an observation in line with the speed accuracy trade-off framework and linked to the effects of response competition on reaction time and response selection (Coles et al., 1995).

Post-Error Responses

Another robust finding from discrete trial studies is that the reaction times in trials that follow error trials are slower than those following correct trials (Rabbitt, 1966b). This finding is similarly well replicated using different discrete trial methodologies (Maier, Yeung, & Steinhauser, 2011; Rabbitt & Rodgers, 1977). A recent study has shown this to be the case even during mental arithmetic tasks. Desmet et al. (2012) asked participants to judge the accuracy of simple multiplication equations and make button presses accordingly. They showed that participants' reaction times were slower after making a wrong judgement compared to after making a correct judgement. Further, they show that this post-error slowing was associated with increased accuracy (81% after correct trials vs. 84% after error trials).

One way to interpret post-error slowing in a performance monitoring perspective is to assert that it serves to improve performance, as predicted by the speed-accuracy trade-off. However, a number of studies have shown post-error slowing is not necessarily associated with post-error accuracy (Hajcak, Mcdonald, & Simons, 2003; Hajcak & Simons, 2008; Rabbitt & Rodgers, 1977). Further, as Rabbitt and Rodgers (1977) observed, the reaction times associated with post-error trials are "far more [slower] than is necessary to return [performance speed] to a 'safe' level" (p.728).

The lack of fit of post-error slowing with the speed-accuracy trade-off framework is further supported by findings of Notebaert and colleagues (Castellar, Kuhn, Fias, & Notebaert, 2010; Notebaert et al., 2009). Notebaert et al. (2009) showed that post-error slowing could be reversed by changing the ratio of error responses to correct responses, such that errors become more frequent than correct responses. Notebaert and colleagues measured post error slowing by subtracting the response times in post-correct trials from post-error trials in a 4-choice reaction time task. They report that in the condition where the participants performance was 75% correct, post-error responses were 25ms slower than post-correct responses. However, when the tasks was made so difficult that the participants performance dropped to 35%, *post-correct* responses became 50ms *slower* than *post-error* responses.

This pattern of results have been replicated with a similar methodology a year later by Castellar et al. (2010). These findings lead to a second account of post-error slowing such that it can be an index of performance monitoring which serves to interrupt behaviour in response to a novel event, and possibly allocate more attentional resources to the ongoing behaviour, without necessarily increasing performance accuracy.

A third account suggests that post-error slowing serves to selective inhibition/suppression of high-conflict responses (Danielmeier & Ullsperger, 2011; K. Ridderinkhof Richard, 2002) following error trials based on the suppression-activation model of K. Ridderinkhof Richard (2002). Available evidence supporting this account comes from behavioural, EEG and fMRI experiments. K. Ridderinkhof Richard (2002) showed in a Simon task that post-error trials were associated with post-error slowing only when the incompatible trials (e.g. left stimulus, right response) were less likely than compatible trials (e.g. left stimulus, left response). If the incompatible trials were the majority, the participants reaction times were already slowed down, and were not associated with slowing following error responses (as compared to following correct responses, Danielmeier & Ullsperger, 2011).

Marco-Pallares, Camara, Munte, and Rodriguez-Fornells (2008) showed that the magnitude of post-error slowing was correlated with the beta band power in an EEG experiment. Increased beta is shown to be associated with motor inhibition (Kuhn et al., 2004), and decreased beta is shown to be associated with faster responses (van Ede, De Lange, Jensen, & Maris, 2011). Evidence gathered in the last decade suggests that the inhibitory beta activity is associated with a network of structures including the subthalamic nucleus, right inferior frontal cortex and pre-SMA (Chevrier & Schachar, 2010; Kuhn et al., 2004; Swann et al., 2009). Further, activity in the right SMA and the dorsal substantia nigra (among other areas) is found to be positively correlated with motor stopping and slowing (Aron et al., 2007) and post-error slowing (Neubert, Mars, Buch, Olivier, & Rushworth, 2010). All these empirical findings add weight to the hypothesis that post-error slowing might be an index of motor inhibition and sub-served by a right hemisphere network.

Pre-Error Responses

Research concerned with performance monitoring has mainly focussed on error and/or post-error performance. Thus, pre-error changes in performance are not as frequently reported as those of error or post-error performance. To our knowledge, the first report of any pre-error changes in performance was published by Smith and Brewer (1995). These authors showed that in a 4 choice reaction time task, responses before an error start (up to about 6 trials before the error) to get faster, and end in a very fast error response. Allain, Carbonnell, Falkenstein, Burle, and Vidal (2004) also reported two experiments where error preceding trials were associated with faster response times than correct preceding trials in flankers and a simple two choice reaction time task. A very similar pattern of performance speed has been reported in a flankers task (H. Eichele et al., 2010), as well as in a modified (4 limb) Stroop task (Gehring & Fencsik, 2001). We are not aware of any reports of pre-error *slowing* in discrete trial tasks, suggesting that when errors are preceded by any change in performance speed, it tends to be an increase in speed.

Conclusions from discrete trial tasks

Overall, findings from a variety of methods using discrete trial tasks suggest that the performance speed before and during an error is in line with a speed accuracy trade-off frame-work. On the other hand, post-error performance is not as easily explained by the speed-accuracy trade-off. Post-error trials are slower than post-correct trials, but they are not always more accurate.

1.1.2 Findings from Continuous Actions

One common property of the studies mentioned above is the discrete manner in which responses are made. Discrete trial tasks are those in which the participant is required to make a button press in response to a target stimulus. This is in contrast to continuous tasks which involve multiple internally generated (e.g. playing a song on the piano or typing a text from memory) or externally generated responses (e.g. playing a song from a score or copy-typing a text).

Error related changes in performance during continuous actions such as typing show certain similarities as well as important contrasts to those in discrete trial tasks. Error as well as post-error responses in continuous actions are found to be slower in continuous actions. The limited evidence on pre-error responses suggest that pre-error responses are also slowed down. Thus there is a stark contrast between the effect of errors on discrete vs. skilled/continuous performance.

These contrasts inform us about how performance is monitored and controlled in more skilled and complex tasks we carry out everyday. We place our focus particularly on typing and piano playing because these tasks involve very similar motor responses to those used by discrete trial tasks. This gives us more confidence in comparing our findings in skilled actions to those in discrete tasks, while enabling generalization of our results to everyday tasks such as writing emails or copying journal article titles to a bibliography for a PhD thesis. In the sections that follow, we describe the behavioural effects of errors on continuous performance.

Error Responses

Unlike those of discrete actions, error actions during continuous performance are typically associated with slower response times than correct ones. For example, Shaffer (1975) showed that in typing, latencies of error key-presses were slower than those of correct key-presses. Similarly, Rabbitt (1978) showed that error key-presses were pressed with less force compared to correct key-presses. Further, those errors which were pressed with less force were more likely to be detected errors (38.67%) than undetected errors (8.9%) (Rabbitt, 1978). When considered together with partial error results, these findings suggest that the slowing down of at least some error key-presses in typing is caused by the intervention of a performance monitoring system, even before the error key is pressed down. Recent piano playing studies further support these results (Herrojo-Ruiz et al., 2009; Maidhof, Rieger, Prinz, & Koelsch, 2009).

An inherent assumption in the claim that error responses are slowed down by performance monitoring is that the single key-presses involved in typing can be interrupted. For many people, typing is a highly practised and automatized action and involves tens of key-presses every second, which are prepared and executed in chunks (i.e. in parallel; Flanders & Soechting, 1992; Rosenbaum, 1991). Whether interruption of single key-presses which are chunked together through hundreds of hours of practice is possible is crucial for the claim that error slowing can be caused by performance monitoring.

Empirical support for the claim that typing of single letters can be interrupted in response to external (e.g. a tone; G. Logan, 1982) as well as internally generated signals (e.g. error detection, F. A. Logan, 1999; Rabbitt, 1978) is indeed documented. G. Logan (1982)'s participants were able to stop typing a word in response to a tone immediately, before pressing a subsequent key in the majority of instances (an exception was the typing of the word 'the'). Similarly, of the 3000 errors made by a skilled typist reported by F. A. Logan (1999), 88% were corrected before a subsequent letter was typed. These studies suggest that typists are able to stop/slow down their error key-presses even before they are completed (as hinted by the findings of Rabbitt (1966a, 1978)).

Slowed error responses have been also observed in piano-playing performance. In a study of performance monitoring, Herrojo-Ruiz et al. (2009), showed that trained piano-players' error key-presses were significantly delayed compared to matched correct key-presses. It is worth noting that these results were present even when the pianists had no visual feedback (see section 2.4 for a more detailed description of Herrojo-Ruiz et al. (2009)'s methodology in comparison to ours). Maidhof et al. (2009) similarly showed that skilled pianists pressed the error note with less force compared to correct notes. To our knowledge these are the only reports of error related changes in piano playing performance.

Error performance of skilled typists and piano-players suggest that the incorrect finger movements are slower than the correct ones. Rabbitt (1978)'s observation that detected errors are more likely to be pressed with less force compared to the undetected errors add weight to the interpretation that what slows participants error responses in typing and piano playing is error detection.

Post-Error Responses

Skilled typists almost always slow down after making a mistake, even when they are not aware of making a mistake (G. Logan & Crump, 2010, 2011; Shaffer, 1975). G. Logan and Crump (2010) gave their participants single words to type and manipulated the visual feedback: For some of the trials that the participant made a typing mistake, the visual feedback falsely showed the word as correctly typed. In some of the trials where the participant typed a word correctly, visual feedback falsely showed the word as typed incorrectly. These authors found that when the participants made an actual typing error, their post-error key-presses slowed down irrespective of the visual feedback. Similarly, when they didn't make a mistake, their post 'error' key-presses did not change, even when the feedback suggested they did make a mistake.

Pre-Error Responses

Pre-error response times in skilled actions have similarly been under-reported compared to those of error and post-error responses. However, the available reports suggest that pre-error responses are slowed down both in typing and piano playing. The pre-error performance of the pianist participants' of Herrojo-Ruiz et al. (2009) was marked by a slowing down of up to 50%: These authors reported that during correct performance, the average time between each key-press (or inter-onset interval, IOI) was 121ms. Strikingly, the average IOI 3 key-presses before the error rose to 190ms. In other words, the 3 key-presses immediately preceding the error were slowed down by 59ms on average. Shaffer's (1975) typists also show a similar but smaller change in performance just before the commission of an error. Shaffer reports a slowing down of between 8ms to 23ms (depending on the type of error) in the median inter-key-press-interval (IKI) for the key-press immediately preceding the error key, compared to that of all key-presses.single

In summary, the limited empirical evidence suggests that the pre-error performance in piano playing and typing is slowed down. The authors who reported these findings saw pre-error slowing effects as caused by the detection of the up-coming error action. Shaffer suggests that "the subject was in difficulty or was distracted" (p.429) while typing the letter before the error. Herrojo-Ruiz et al. (2009) suggested that pre-error slowing 3 keys before the error can be an index of early error detection enabled by the internal forward models (Wolpert & Miall, 1996). In other words, these authors suggest that the pre-error slowing is possibly an effect of the detection of the upcoming error.

Conclusions from Continuous Actions

In summary, there is convincing evidence that the errors and post-error responses in skilled, continuous actions are slowed down. There is some evidence that the responses preceding the error are slowed down, but this is not a finding as well replicated as error and post-error slowing.

1.1.3 Conclusions from Behavioural Effects of Errors

Overall, certain behavioural changes in performance speed are shown to be useful parameters in studying performance monitoring processes. Post-error actions in most tasks; and error and pre-error performance in continuous tasks have been shown to be slowed down by errors and possibly by performance monitoring. This body of research based on behavioural effects has a number important implications for the study of human cognition: 1) It suggests that our nervous system is sensitive to the upcoming outcomes of its actions, and uses this internal feedback to adjust ongoing behaviour. This provides some justification in using the term *performance monitoring* to refer to the neural processes involved in altering the ongoing behaviour in response to performance errors. 2) This literature provides testable predictions about expected behavioural effects of errors on performance depending on the nature of the task. Importantly, if these behavioural changes are caused by an internal monitoring process, they must be *preceded* by neural activity underlying this monitoring process.

The section that follows outlines the electro-physiological findings from the human studies of performance monitoring and error detection.

1.2 Electro-physiological Literature

In the following sections, we focus on findings from EEG studies of performance monitoring. These include ERP findings such as the ERN (Gehring et al., 1993) also called Ne (Falkenstein et al., 1991), as well as changes in EEG oscillations. Even though we mention findings from magneto-encephalogram (MEG) and functional magnetic resonance imaging studies (fMRI), our main focus is on the EEG literature as this is the neuro-imaging method used in the studies reported in the subsequent chapters.

1.2.1 EEG studies of Performance Monitoring

More than two decades of EEG research has established a robust error related ERP component. The ERN has been used as an index of performance monitoring and error detection in a large variety of behavioural tasks, in hundreds of studies since the early 1990's. More recently, increased power and synchrony in theta (4-8Hz) oscillations have been shown to underlie the ERN. However, the exact nature of the interaction between these electro-physiological effects with other cognitive functions and adaptive changes in behaviour are still not as well documented.

1.2.2 Error Related Negativity

The ERN is an ERP that follows error responses. The defining features of ERN, as of any ERP, are its polarity, latency and topography. As its name suggest, it is a *negative* going ERP component, which peaks about 50-100ms following an error response, and is best recorded at the fronto-medial electrodes such as the FCz (see figure 1.1).

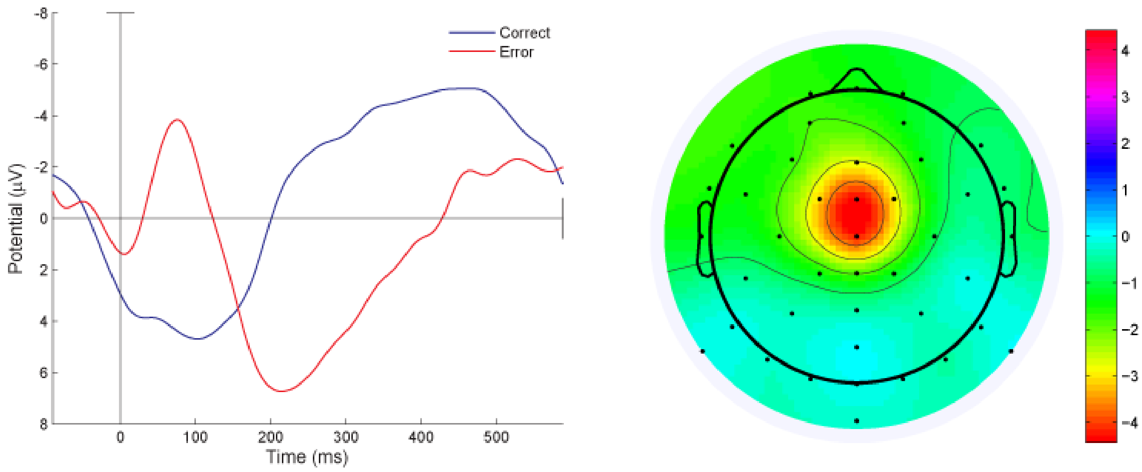


Figure 1.1: The ERN; latency, polarity and topography. Figure reproduced with permission from Willoughby (2012)

ERN After Correct Responses

Some researchers argue that ERN is an ERP component not specific to error responses (Vidal, Burle, Bonnet, Grapperon, & Hasbroucq, 2003; Vidal et al., 2000). These researchers pointed out that the electrical activity after a correct key-press also becomes more negative compared to before. The latency and topography of this correct-following negativity is almost identical to the ERN. However, the amplitude of the post-correct ERP is considerably smaller than that of ERN.

We would like to outline from the beginning that our approach to defining ERPs in the current thesis is slightly different from that used by Vidal and colleagues. Our definition puts greater emphasis on the differences between an event of interest (i.e. error key-press) and another baseline event (i.e. matched correct key-press), than on the differences between before and after the same event. Thus our definition of the ERN is based on the difference between the electrical changes that follow the error responses and those that follow matched correct responses (e.g. a correctly typed 'a' vs. an incorrectly typed 'a').

ERN Latency, Amplitude, Neural Generator

Most of the EEG studies of performance monitoring have used discrete trial tasks similar to the ones mentioned in the previous section on behavioural results. These studies have shown that even though the amplitude and the latency of the ERN peak is sensitive to the methodology and task demands, its amplitude is always negative with respect to baseline, latency of its peak is 50 to 100ms after the response onset (Falkenstein et al., 1991; Gehring et al., 1993; van Veen & Carter, 2002), and is typically best recorded at mid-line fronto-central scalp locations (e.g. the electrode FCz in the 10-20 electrode system). This lead to the hypothesis that the neural generator of ERN is within the anterior cingulate cortex (ACC), a finding supported by fMRI, (Carter et al., 1998; Fiehler, Ullsperger, & von Cramon, 2004; Ullsperger & von Cramon, 2001; van Veen, Cohen, Botvinick, Stenger, & Carter, 2001) and MEG studies (Miltner et al. (2003), see Gehring, Liu, Orr, and Carp (2012) for a comprehensive review chapter on ERN including its neural substrates).

It is important to note here however that the involvement of ACC is not limited to performance monitoring and includes cognitive, emotional, motor, nociceptive and visuo-spatial functions (see Bush, Luu, and Posner (2000) for a review of the functions sub-served by the ACC as well as an overview of the functional connections between parts of ACC and other sections of the brain).

ERN in Skilled Continuous Actions

To our knowledge, there are only two reports of an ERN in skilled continuous tasks (Herrojo-Ruiz et al., 2009; Maidhof et al., 2009). In both of these studies skilled piano players were recruited to play pieces of classical music as the EEG was recorded. Herrojo-Ruiz et al. (2009) found that between 20-70ms *prior* to the error responses, a negative peak was observed at the FCz electrode. Herrojo-Ruiz et al. (2009) called this negativity pre-ERN. These findings were almost identical with those of Maidhof et al. (2009). In a more recent paper Herrojo-Ruiz, Strbing, Jabusch, and Altenmuller (2011) showed that the error related changes were not only reflected in the EEG record but also in the time-frequency decomposition of the error related EEG (see section 1.2.4 for a short discussion of error related changes in EEG oscillations).

1.2.3 ERN and Error Awareness

The observation of a change in the ACC activity in response to errors suggests that the ERN might be an electro-physiological marker of error detection. However, the observation of ERN after undetected errors poses difficulties for this interpretation. A number of computational models have been developed to answer the question of what neural process the ERN is a manifestation of.

The most prominent computational models are the conflict monitoring model of Botvinick, Braver, Barch, Carter, and Cohen (2001) suggesting a direct link between the amount of response conflict between current response choices and the ACC activity (but also see Burle et al., 2008). Another prominent computational model is that developed by Holroyd and Coles (2002), which suggests that the ERN is a negative reinforcement signal used by the ACC to modify on-going behaviour.

See Alexander and Brown (2010) for a review and synthesis of competing models of ACC functions.

None of the models mentioned above draw a direct link between ERN and conscious error awareness. This is because many EEG and fMRI studies show that the ERN and ACC activation are present even after errors which the participant is not aware of making (Ehlis, Herrmann, Bernhard, & Fallgatter, 2005; Endrass, Franke, & Kathmann, 2005; Endrass, Reuter, & Kathmann, 2007; Hester, Foxe, Molholm, Shpaner, & Garavana, 2005; OConnell et al., 2007; Ursu, Clark, Aizenstein, Stenger, & Carter, 2009). However, some of these studies showed another component of the error related EEG record which was indeed sensitive to conscious awareness of the error commission, namely the Pe.

Error Related Positivity

The Pe follows error onset by about 200 - 400ms and its amplitude is modulated by the awareness of error commission (Endrass et al., 2005, 2007; OConnell et al., 2007). As such, even though the ERN amplitude is no different after detected vs. undetected errors, the Pe amplitude is considerably larger after detected errors than undetected errors. These findings, together with its later onset, suggests that the Pe might be involved in the processing of the error response at a higher, conscious level.

1.2.4 Errors and Brain Oscillations

In addition to their effect on ERPs, a growing literature suggests that errors also affect brain oscillations of certain frequencies. Observation of theta activity during behavioural tasks is not new. For example, Laukka, Jarvilehto, Alexandrov, and Lindqvist (1995) cite a study conducted in 1950 as the first report of EEG theta activity during a problem solving task. Laukka et al. (1995) themselves showed increased theta power during learning in a simulated driving task. These authors showed that the theta activity was highest during the parts of the task where the participants had to choose what road to take (out of 3 or 4 alternatives) in response to an imperative stimulus. Similar findings relating theta activity to behavioural as well as cognitive performance have followed since then (Campagne, Pebayle, and Muzet 2004; Jensen and Tesche 2002; Mazaheri, Nieuwenhuis, van Dijk, and Jensen 2009, but also see Mizuki, Tanaka, Isozaki, Nishijima, and Inanaga 1980).

More importantly, when compared to correct responses, incorrect responses have been shown to be associated with much stronger time-locked increases in power and synchrony of theta band oscillations (Cavanagh, Frank, Klein, & Allen, 2010; Luu, Tucker, Derryberry, Reed, & Poulsen, 2003; Luu, Tucker, & Makeig, 2004; Trujillo & Allen, 2007). The literature on the effects of error responses on theta oscillations during discrete trial and continuous tasks is covered in detail in the introduction section of chapter 5. In order to avoid duplication, we see it sufficient at this point to state that the errors of performance have been associated with increases in power and synchronization of oscillations in the theta frequency band.

1.3 Gaps in the Literature

At any given point in time, complex and dynamic neural systems are active to varying degrees in the brain. Singling out one hypothetical cognitive mechanism, calling it 'performance monitoring', and attempting to study it has obvious theoretical as well as practical difficulties. One such difficulty is identifying the parameters associated with it. However, thanks to decades of research, now we can stand on the shoulders of giants and point to certain behavioural and electro-physiological parameters sensitive to errors of performance, such as post-error slowing, the ERN and theta oscillations. However, how these parameters are related to each other and interact in the context of performance monitoring during ecological tasks like typing to improve or adjust performance are still not well-understood.

1.3.1 Gaps in Literature and Our Proposed Contributions

Despite the robust nature of electro-physiological markers of performance monitoring, there is no consensus about how these processes are translated into behavioural adjustments. For example, the observation of ERN after undetected errors, or the lack of correlation between the error related behavioural (i.e. post-error slowing) and electro-physiological (ERN or Pe amplitude) parameters make it difficult to draw conclusions about the exact function served by these EEG effects.

Revisiting the second point raised in section 1.1.3, we are not aware of any empirical work which suggests that the electro-physiological indices of error detection necessarily precede behavioural effects of error detection in skilled and continuous actions: Herrojo-Ruiz et al. (2009) and Herrojo-Ruiz et al. (2011)'s results show the errors lead to changes in EEG (i.e. the pre-ERN and theta changes) about 100ms before an error whereas the changes in behaviour (pre-error slowing) start more than 500ms before the error. Rabbitt (1978) and Shaffer (1975)'s observations of slowed and weakened error key-presses during typing have not been studied using EEG so there is no empirical support for the claim that the error slowing observed in typing is preceded by the neural markers of error detection.

Thus one area we aim to contribute with the current studies is the following up of the robust behavioural effect reported by Rabbitt and Shaffer more than 40 years ago using EEG, extending the time of interest to before error onset periods.

1.3.2 Interplay between Performance Monitoring, Cognitive Control and Adaptive Behavioural Change

Observing the detrimental effects of lateral pre-frontal lesions on both the behavioural and EEG indices of performance monitoring, Gehring and Knight (2000) suggested that the medial parts of the brain (such as the ACC) must be interacting with the lateral pre-frontal to areas to enable changes in the ongoing behaviour. Recent studies such as those by Cavanagh, Cohen, and Allen (2009) and H. Eichele et al. (2010) revealed promising functional connections between these error related changes in the medial parts of the brain and more dorso-lateral parts of the pre-frontal areas, which have traditionally been associated with higher level cognitive functions such as attention. Further, these functional interactions were found to predict error related behavioural changes. To add further weight on the importance of investigating neural connectivity, Cohen (2011) showed that the stronger the connections (i.e. white matter connectivity as assessed by diffusion-tensor imaging) between the error sensitive medial areas (such as the ACC) and certain other regions in the brain, i) the stronger the error induced theta power recorded over fronto-medial electrodes; and ii) the better synchronized the theta oscillations are between fronto-medial and other electrodes.

These recent findings draw our attention to the importance of considering the performance monitoring system as a part of a larger, interactive system to better understand their involvement in adaptive behavioural adjustments. However, even when the interactions between performance monitoring and other neural systems are taken into account, electro-physiological and behavioural effects of errors are not consistently found to be correlated (Cohen 2011; Gehring and Fencsik 2001; Gehring et al. 2000; Grundler, Cavanagh, Figueroa, Frank, and Allen 2009; Hajcak et al. 2003 but also see Cavanagh et al. 2009; Gehring et al. 1993; Huster et al. 2011; R. K. Ridderinkhof, Nieuwenhuis, and Bashore 2003). We believe using an ecological task with a natural error signalling response can potentially inform how and if the EEG and behavioural measures interact in the context of performance monitoring.

1.3.3 Current Studies

We believe the main target of cognitive psychology should be to understand how cognitive abilities such as performance monitoring enable adaptive changes to the behaviour of the agent as a whole. Knowing what parameters can be used to index performance monitoring is necessary but not sufficient to reach this target. Unless we can find a way to improve people's abilities (be it tea making or driving) in the real world, our job is incomplete.

Accordingly, a primary aim of the studies reported here is to investigate electro-physiological measures (namely power and synchrony of theta oscillations) not only within the medial sources of electro-physiological activity, but also in connection with other oscillatory sources (chapter 5), in addition to classic measures such as the ERN and Pe (chapter 4) and post-error slowing (chapter 3) with a particular interest in any correlations between these measures (chapters 4 and 5). A novel contribution of the studies introduced here is the fact that these indices of performance monitoring were recorded during a complex and ecological task, which the readers of these lines spend tens of hours performing on a weekly basis. Once the interaction between these parameters related to errors and performance monitoring are better understood, we can begin looking for ways to improve people's performance in everyday and skilled/specialized tasks.

Chapter 2

Methodology

2.1 Introduction

The aim of this chapter is to provide a detailed description of the methodology used in our experiments. In addition, technical and analytical problems faced and the approaches used to overcome them are described. In many ways, our methodology is based on previous research. For example, the typing task we used in both experiments is almost identical to that used by Rabbitt (1978), except that it is computerized, and that we recorded typing speed as well as accuracy. Similarly, our second, EEG experiment is very similar in methodology to that conducted by Herrojo-Ruiz et al. (2009), except i) we used typing rather than piano-playing as the behavioural task, and ii) we used a different approach in the analysis (i.e. ICA) of our EEG data. A more in-depth description of our methods and how they compare to previous research is discussed in the following sections.

Participants Participants who took part in our experiments were skilled touch-typist. However, in a pilot study not presented here, we first recruited undergraduate students, who could type very fast (up to 80 words per minute). However, we soon realized that in the absence of visual feedback, these participants could not type accurately. More specifically, almost all of these participants shifted their hand position in the middle of a trial at some point in the experiment. This meant that the participants were typing assuming their hands were in the right place, when they were not. As a result, these participants ended up typing letters which were physically adjacent to the correct ones on the keyboard (e.g. typing 'upi' instead of 'you').

This left us with no data from the pilot study, but also taught us a lesson: Recruit participants who are trained or can reference their hands back to where they need to be without having to look at the keyboard. After this point, our criteria for recruiting participants became more strict. When sending out invitation emails, we asked whether people were familiar with the 10 finger typing system, or if they had formal training through a course, or informal training using any typing software. It is important to note that some skilled typists also shifted their hands during the experiment, but they were quick to fix it because they knew how to reference their fingers on the keyboard without looking at them.

2.2 Behavioural Experiment Design

The behavioural task used in both studies reported in this thesis was typing, and was almost identical to the methodology used by Rabbitt (1978) except for a number of key points. One key difference was the use of a computer keyboard to record key-presses in the current experiments, as opposed to a type-writer. Using a computer enabled us to easily record the time of each key-press, which is not as straightforward with a type-writer.

An important issue in the studies of cognitive control and error detection is error signalling. Error signalling by the participant allows the external observer (i.e. the researcher) to decide whether the participant realized that they made an error or not. To signal subjective detection of an error, the participants of Rabbitt (1978) were instructed to press the star (*) key immediately after making a mistake. This is not a natural response to typing errors, and definitely not a part of the skills of a typist. In the current experiments, participants were asked to correct any mistakes they made by pressing backspace, as they would when typing naturally.

The main focus of Rabbitt (1978) was to test whether the typists could detect their mistakes in the absence of visual feedback and if so, how fast. He showed the typists could detect almost all of their errors, most of them even before they typed a subsequent post-error key. Our main focus was to test whether the robust effects of errors on performance, such as post-error slowing, would extend to the period before the error. Specifically, a question addressed was whether a recently reported pre-error slowing effect (Herrojo-Ruiz et al., 2009) generalizes to typing. Thus in addition to the ability to detect and correct ones mistakes, we investigated the time at which errors and error detection started to have an effect on ongoing behaviour.

As mentioned in section 3.2.2, the participants in our experiment typed the text presented with no visual feedback. The reasoning behind this methodological choice was to replicate and extend the findings of Rabbitt (1978) (and also Herrojo-Ruiz et al. (2009), see below) in two respects: 1) Error detection is possible in the absence of visual feedback, and more importantly 2) electro-physiological markers (e.g. the onset of ERP and theta oscillation effects) of error detection do precede the behavioural markers of error effects (e.g. error or post-error slowing). This temporal relationship is a fundamental prerequisite for the claim that it is error detection that leads to error slowing. This temporal relationship has not been empirically shown in the literature to our knowledge. Further, G. Logan and Crump (2010) showed that the outer loop relies on the feedback from the monitor whereas the inner loop relies on the feedback from fingers. By eliminating feedback from the monitor investigated the relationship between inner and out loops and the feedback from the monitor.

On the other hand one obvious drawback of eliminating visual feedback is that it compromises the ecological validity of task at hand. Most people type with visual feedback from the monitor, and their hands in their sight. However, we believe that the benefits of comparing our results to previous work by removing the visual feedback from the task is an acceptable compromise in the long run, and further studies can be conducted to see if the empirical results we report here will generalize to typing with full visual feedback.

2.2.1 Recording of Typing Performance

Key-presses and their timing were recorded using ©MATLAB running on Microsoft Windows XP. We wrote a custom MATLAB script which checked the status of the keyboard every 2ms and recorded i) the key that was pressed and ii) the time of the key-press. Key-press times were recorded such that the very first key-press had a time of zero, and the time of the subsequent key-presses were the time elapsed since time zero. The absolute time for each key-press was transformed into IKIs such that the IKI for the current key-press would be the time elapsed between the pressing of the previous key-press and the current one. This way, changes in typing performance could be assessed at the key-press level.

We investigated the effect of errors within words. For example, if the 5th word contained an error letter, the analysis of pre- and post-error slowing was confined to the 5th word only, and not extended to the 4th (for pre-error analyses) nor the 6th word (for post-error analysis).

2.2.2 Coding of Errors

We used a semi-blind procedure to identify and classify the error words. An algorithm was developed to automatically tag words which were not typed correctly using MATLAB. These errors were then manually coded such that the following information was recorded for each error: Error letter position within the word, word length, type of error, whether the error was followed by backspace or not, whether the error was a genuine error or a false alarm (i.e. correct key-press followed by a backspace).

False alarms were not included in the analysis, and all error types, except for omissions and insertions were collapsed because we found that they were not statistically different from each other in terms of IKI. Justification for excluding omissions and insertions is provided in the methodology section of chapter 3, but very briefly, omitted key-presses (i.e. omissions) do not have an IKI, and insertions are simultaneous key-presses, caused by a single finger pressing down more than one key simultaneously.

2.3 Behavioural Data Analyses

2.3.1 Individual Differences in Typing Speed

Average IKI varies from one person to another. An IKI of 300ms might be among the slowest IKIs for one typists, but be close to the average IKI for another one. Thus, the average IKI for key-presses would vary not only with the accuracy of the key-presses (as shown previously by Rabbitt (1966b, 1978)), but also with individual differences in typing speed. In order to minimize the amount of variance contributed to IKIs by individual differences, we calculated an error slowing measure for each error key-press within participants. Thus, rather than using the average of raw IKIs for each error key-press and comparing it to that for correct key-presses, we simply calculated the error slowing value for *each error key-press* by subtracting from it the average IKI for correct key-presses (see the section 2.3.3).

This way we evaluated the effect of error on typing speed of each participant, and used this parameter in all of our statistical tests. Consequently, a positive effect (i.e. a positive error slowing) indicates that error key-presses were slower than correct key-presses, and vice a versa.

To provide a reference to judge the typing speed of participants who took part our study we provide here the average typing speed of our participants, and those of a number of other typing studies. Average speed of our participants was 67wpm (SD = 6.39) for the first experiment (chapter 3), and 70wpm (SD = 13.32) for the second experiment (chapters 4 and 5). The typing speed of the *expert typists* of Gentner (1983) ranged from 61 to 90wpm. The range of typing speed of participants in the study of Inhoff (1991) was 51-116wpm. Typing speed of typist participants of G. Logan (1982) ranged from 47-79wpm. The single participant who took part in the study of Shaffer (1975) could type more than 100wpm. Based on these data, we believe the skill level of our participants was comparable to those considered to be skilled typists in the literature.

2.3.2 Non-parametric Tests for Error Slowing Analyses

Applying parametric tests such as t-tests, or using measures such as standard deviation which are based on the assumption that the data at hand are normally distributed, especially when the sample size at hand is not large, has obvious drawbacks. Even though the number of error key-presses recorded was large in both experiments reported here, the same was not true for the number of pre-error key-presses, particularly for Experiment 1, where participants typed only 20 sentences. This is because not all errors were preceded by correct letters (e.g. first letters of words have no pre-error letters).

We used non-parametric tests such as Wilcoxon's signed ranks test for comparing error slowing to 0. Wilcoxon's signed ranks test is commonly viewed as a non-parametric analogue for paired t-test (Howell, 2002). The confidence intervals around the average error slowing values were constructed using the bootstrap technique which is basically a sampling with replacement technique and does not require the data to be symmetric (Howell, 2002). Further, the boundaries of the confidence intervals constructed by the bootstrap technique reflects the shape of the distribution: If the distribution of the data is positively skewed, the upper bound of the confidence interval will be further away from the average than the lower bound.

2.3.3 Non-Error Factors affecting the IKI

We found, in line with previous research that key-presses in longer words had larger IKIs than those in shorter words (Rosenbaum, 1991; Shaffer & Hardwick, 1969). More importantly, key-presses in longer words were more likely to be errors on average than shorter words. For example the 2^{nd} letter in a 7 letter word on average had a slower IKI and a greater error likelihood than that in a 3 letter word. Thus, word length in our experiment was a confounding variable affecting the typing speed. We found a similar effect of letter position on typing speed, such that the first letters had longer IKIs than the subsequent letters in words matched for word length, and the subsequent IKIs increased with increasing letter position.

In order to eliminate the confounding effects of word length and letter position, we calculated the error slowing in the following way. First, for each participant, we created a look-up table containing average IKIs for correct key-presses, such that the rows represented the word length, and the columns represented the letter position. For example, the number in the 4^{th} row and 2^{nd} column of this table would contain the average IKI of the all 2^{nd} letters in 4 letter words. When calculating the error slowing value associated with each error, numbers from these tables were used. As such, the average IKI for correct key-presses *matching the letter position and word length* of the error key-press was used to calculate error slowing value, such that:

$$\text{Error Slowing} = \text{IKI}_{\text{ }Error\ Keypress} - \text{IKI}_{Avg.\ Matched\ Correct\ Key-presses}$$

2.4 EEG Methods

The design of our experiment was also similar to that used by Herrojo-Ruiz et al. (2009). These authors asked skilled piano-players to play previously practised pieces while recording EEG from 35 electrodes. The pianists played under 3 different conditions, one of which was identical to ours. In this condition, the participants could not see their hands or hear the outcome of their performance. The main focus of Herrojo-Ruiz et al. (2009) was the effect of errors on behavioural performance and also on the electro-physiological recordings, in a skilled and continuous task. From this perspective these authors' goals and methods were similar to ours.

One of the key differences between our and Herrojo-Ruiz et al. (2009) study is that between typing and piano playing. Herrojo-Ruiz et al. (2009) also used IKI (but called it IOI for inter-onset interval) to assess performance. We believe that typing is a more sensitive task than piano playing for detecting changes in performance speed. This is because there is a clear distinction between correct and incorrect key-presses in typing and it only depends on the letter typed. On the other hand, in piano playing *the timing* and the strength of the key-press as well as the note played are important for the accuracy of performance. The importance of this distinction between piano playing and typing becomes obvious when the speed of error key-presses are compared to speed of correct key-presses, and this issue is discussed in depth in the discussion section of chapter 3.

A second difference is that Herrojo-Ruiz et al. (2009) conducted all of their analyses on electrode data, cleaned from artefacts using ICA (see section 2.4.3 for a brief description). In other words, after applying the ICA, these authors identified and removed the independent components (IC) representing the muscle artefacts from the raw electrode data. Then, they conducted all of their statistical analyses on the remaining ICA-pruned electrode data. On the other hand, we extended the use of ICA into our statistical analysis. More specifically, for each participant, we identified the ICs of interest, and used these ICs rather than the cleaned or raw electrode data. A more detailed description of our use of ICA in the analysis is provided in section 2.4.3.

The remaining sections provide a description of the methodology used in the EEG experiment, technical and analytical challenges faced, and approaches used to solve them.

2.4.1 Procedure

All testing was conducted in a well lit room in the Department of Psychology, University of Sheffield. When the participants arrived, they were given a step by step description of the experimental procedure, and then asked to sign an informed consent form. After signing the consent form the participants were seated in front of the testing computer and asked to adjust their distance from the monitor and the keyboard such that they could type comfortably. The recording computer was connected to the Biosemi USB Box, which received signals from the EEG amplifier as well as the testing computer.

When the participant was ready, the EEG cap was placed on her head, and electrolyte gel was injected to provide the electrical connectivity between her scalp and the electrodes. In order for the participants to get used to the feel of the keyboard, they were asked to type whatever they wanted (e.g. check emails or play on-line typing games on the website `www.typingtest.co`) during the EEG preparation. Once all the electrodes were connected to the amplifier and the recording software was ready, the participants were asked to watch their own on-line EEG and make eye movements, clench their jaws, and make other movements which typically result in EEG artefacts. Then they were explained that these movements would contaminate the EEG record, and thus they should refrain from such movements while they are typing. Since the typing task was self paced, they were asked to keep these movements to between-sentence-periods.

2.4.2 Time-Stamping of EEG and Key-press Data

One of the most important advantages of using EEG as a neuro-imaging technique is its high temporal resolution. However, in our methodology, in order to exploit this advantage maximally, the recording of EEG data should be well synchronized with the recording of the key-press data.

As a first step, we acquired a high speed keyboard (©DirectIN PCB v2010 from Empirisoft, `http://www.empirisoft.com/directinkb.aspx`), the internal wiring of which is modified such that the key position information is collected from each button within 1 ms, and the testing computer is instructed to poll the keyboard every 1 ms. We used this key-board in the experiments reported. This way we ensured that the delay between a participant pressing down a key and the computer recording is minimized.

Second, rather than recording the EEG (through Actiview Software on the recording computer) and key-press data (through MATLAB on the testing computer) separately and combining them offline, we recorded them together using the same hardware. We achieved this by changing the custom MATLAB code such that the testing computer sent the key-press information (i.e. the ASCII code for the letter pressed) to the Actiview USB Box `http://www.biosemi.com/` online, as the participant typed. Time stamping of EEG signals is accomplished by the Actiview USB Box. This way, even before the Actiview software on the recording computer received the EEG signals, the timing of key-presses was incorporated with EEG data.

We applied the following check to make sure that the latencies of key-press events as recorded by the Actiview software on the testing computer and MATLAB on the recording computer were synchronized: We plotted the difference between the IKIs as recorded by MATLAB and those recorded by the EEG software. We found that the difference between IKIs recorded by the equipments was negligible: The mean delay during recording from a participant was 0.11ms (SD = 0.11ms). Given that the median IKI for our typists was 155.81ms (SD = 19.43), the fastest typing at an IKI of 123.92ms, the lag of less than 0.2ms for the transmission of signals was considered acceptable.

2.4.3 Independent Component Analysis

Independent Component Analysis is a statistical blind source separation technique. Exploration of the mathematical details of ICA is beyond the scope of this thesis, and it will be described only as far as is useful to understand its effects on EEG data (see Groppe, Makeig, and Kutas (2008) for a review of the use of ICA in EEG analysis).

One aim of EEG research is to infer neurological function from the electrical changes recorded over the scalp. One assumption inherent in this statement is that the electrical changes over the scalp are caused by neural processes. However, we know that i) the electrical changes recorded at the scalp are not affected solely by neural function and ii) multiple distinct neural processes overlap in time and space. As such, EEG recorded over the scalp at any moment in time is a linear mix of multiple neural processes (each of which will be activated to a different extent at different times), and other sources of electrical activity such as the muscular activity from the eyes or the scalp itself (which have much stronger voltage amplitudes than those of neural activations).

Under optimal conditions, the input to the ICA are the linearly mixed signals, and outputs of the ICA are i) the independent signals (or independent components), which make up the observed/mixed signals when linearly summed together and ii) the mixing matrix which describes the relationship between the mixed input signals and the latent independent components (IC). It is easier to explain this in a mathematical manner:

$$\mathbf{x}(\mathrm{t}) = \mathbf{A}\mathbf{u}(\mathrm{t})$$

where $\mathbf{u}$ is an n-dimensional vector of independent sources whose activity at time t is linearly mixed via the transformation matrix $\mathbf{A}$, to create the n-dimensional observations $\mathbf{x}$. Using the same terminology as the previous paragraph, the input to ICA in the context of EEG is the n-dimensional $\mathbf{x}$ (i.e. the recordings from n electrodes), and the output are the mixing matrix $\mathbf{A}$ and n independent sources $\mathbf{u}$. It is important to note here that ICA is blind to electrode location.

Once EEG data have been decomposed into the ICs, there are a number of ways one can proceed with the data analysis. Many researchers use ICA to identify muscle artefacts and remove them from the neural activity. This is in a way "subtracting" the muscle activity on EEG from neural activity. However, as mentioned in the second paragraph of this section, EEG is a mix of muscle as well as *multiple* neural generators of electrical activity. Subtracting EEG due to muscle activity from neural activity solves only one part of the problem, and leaves us with the linearly mixed signals from multiple neural generators: De-mixing muscle activity from neural activity leaves the temporally overlapping neural activity mixed. In the grand average ERPs, the activations of neural generators which are not time-locked to the event of interest are theoretically averaged out. However, in single trial analyses, this would decrease the signal to noise ratio and inflate the rate of type II errors.

Justification for the Use of ICA in the Current Analysis

We decided to use the ICA not only to remove blink and muscle artefacts, but also to further decompose the EEG data into ICs representing temporally independent time series of activations, which theoretically represent different neurological processes. One might object to the selection of the use of ICA simply because the ERN and the underlying theta burst are very robust and well-replicated EEG effects even in the absence of ICA or other spatial filtering methods such as the current-source density transforms (CSD, Kayser and Tenke (2006)).

Two considerations have been influential in our decision to use ICA. The first consideration was related to the behavioural task used. Typing is a dynamic task where the participants continuously read, make eye movements, and prepare and execute coordinated bi-manual finger movements. Many of these functions are served by cortical structures, which may or may not be spatially close the generator of error related responses (i.e. the ACC). Further, muscular activity from the eyes as the participant reads would constitute large amplitude EEG signals which would not be possible to remove simply by cutting out as artefacts because they would be present continuously in the EEG record. From this perspective, ICA seemed to be an ideal solution to separate error related EEG activations from reading and typing related motor activations.

Second, we evaluated previous performance monitoring investigations using ICA as an analytical approach. A number of publications already exist which show that ICs representing the ERN (i.e. error related IC, or ErIC) are very robust signals readily isolated by ICA (e.g. Debener et al., 2005; Gentsch, Ullsperger, & Ullsperger, 2009; Hoffmann & Falkenstein, 2010; Murphy, Robertson, Allen, Hester, & O'Connell, 2012; Roger, Benar, Vidal, Hasbroucq, & Burle, 2010; Wessel, Danielmeier, Morton, & Ullsperger, 2012). In fact, in their review of ICA as a means to study ERPs, Groppe et al. (2008), suggest that "probably the most convincing extraction of a source of a cognitive ERP to date is [...] the 'error-related negativity' (pp. 11)".

For these reasons we decided to go ahead with using ICA to separate error related activations from non-error related activations in EEG. Below, we give more specific examples from the data which warranted and explicitly showed the effect of decomposing the data by way of comparison to electrode data.

Separating Error Related Activity from Motor Activations Continuous copy-typing involves production of very fast and well coordinated bi-manual finger movements. The ERN has been shown to be generated in the ACC which is in close proximity to motor areas extending well into the medial wall of the hemispheres. The supplementary motor area in particular is an essential structure in the preparation and execution of bi-manually coordinated finger movements (Sadato, Yonekura, Waki, Yamada, & Ishii, 1997) and skilled actions like piano playing (Gerloff, Corwell, Chen, Hallett, & Cohen, 1997).

Because of the close proximity of the motor and pre-motor areas to the ACC, the electrical data recorded at fronto-medial electrodes are likely to be a mix of activations generated by these two (and possibly more) distinct electrical sources. However, because the motor areas would be activated during the execution of all key-presses, whereas ACC activity has been shown to be more specific and sensitive to error responses, they are expected to have different temporal contingencies. It is exactly problems like this that ICA is designed to solve (Delorme & Makeig, 2004; Groppe et al., 2008; Onton & Makeig, 2006).

Visual inspection of participant data revealed that the trial-by-trial pattern of electrical activity in ErIC closely reflected that of fronto-medial electrodes (see figure 2.1), in terms of latency and topography. However, in line with the expectation of overlapping activations of multiple sources at the time of error, we found that the amplitude of the voltage changes were considerably smaller in the ErIC compared to those in fronto-medial electrodes.

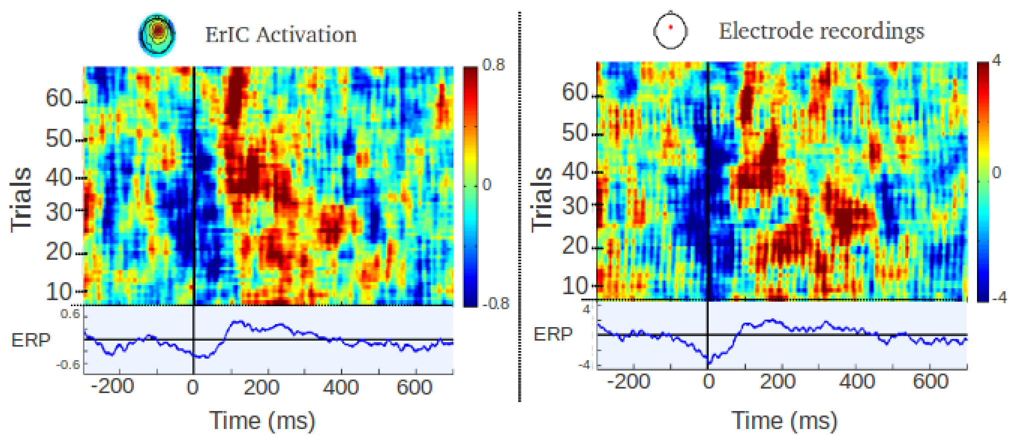


Figure 2.1: Magnitude of electrical potentials (in micro-volts, as represented by the colour scale) recorded by the electrode (right) vs. that after the application of ICA (left). The considerable difference used in the colour scales to plot these figures clearly shows the effect of ICA on the electrical activity recorded. Data from a single participant.

A close inspection of the data showed that this difference was uniform over all time-points and events (i.e. error and matched correct key-presses). We've confirmed that this was the case by visually comparing the activity of ErIC to the fronto-medial electrodes in single trials as well as in average ERPs for each participant.

Cleaning of Other Error Related Activations During visual inspection of EEG data, we found that some participants blinked after the commission of errors. We realized at this point that errors may trigger behavioural and arousal reactions as well as blinks, which may at times temporally overlap with the performance monitoring processes we were interested in. In fact, the error positivity (i.e. Pe), has been shown by previous research to temporally and spatially overlap with perceptual components of the ERP (e.g. P300, Leuthold and Sommer 1999; Nieuwenhuis et al. 2001).

Source Localization Using ICs instead of electrode data theoretically results in more accurate scalp maps for a given EEG effect. This in turn leads to more accurate models for finding the most likely location of the generator of that EEG effect. This issue is discussed in more depth in the clustering section (section 2.4.4). Briefly, this is because scalp map for an IC representing the EEG effect of interest is created *after* it is isolated from other EEG effects which might be active at the same time and in close proximity.

Considering all the points raised above, we believe that the choice of ICA to extract error related EEG activations was well-justified. Accordingly, all of the analyses and results reported in chapters 4 and 5 are based on ICs, and not on raw EEG data from individual electrodes.

2.4.4 Cluster Analysis

Implications of Clustering for Across Participant Analysis

Using ICs instead of electrode data brings many advantages to EEG analysis, but it also complicates the analysis specially at the across participant level. One such complexity is related to the selection of equivalent ICs to cluster together.

Clustering of EEG data from electrodes across participants is straight forward, as data from each electrode is assumed to be comparable with those from corresponding electrodes for all participants (e.g. FCz on participant 1 and FCz on participant 2, etc.).

With ICs however, the selection of the IC that represent the same activity across participants is more complicated, and necessitates some form of selection for each participant. The selection may be based on a number of different approaches, including subjective methods like visual inspection and blind methods like automated algorithms.

We believe the cost of this added level of complexity is well compensated for by its benefits. This is particularly the case when the analysis is well informed by previous research, and one has *a priori* expectations as to when and where the EEG signal of interest will appear. For example, we knew that the ERN is a negativity associated with errors, generated in the area around ACC and most likely to be detected by the fronto-medial electrodes. Thus, we could look for ICs with these properties when clustering them together.

It is generally assumed that the electrical potentials recorded by the same electrode (say FCz) on different participants will represent equivalent processes. However, the electrical potentials recorded at the scalp are sensitive to the orientation of the dipole generated by the electrical source. Thus, a source with a fixed 3D location within the brain might generate different scalp maps depending on the orientation of its dipole (which itself is affected by the structural properties of the neural generator such as the sulci and the gyri), which may be different for every participant. So using 3D locations of their generators is a more robust approach than using scalp maps or electrode positions for clustering equivalent processes.

Because ICA decomposes EEG data from each participant into temporally independent components and *then* calculates the scalp map, it is automatically adapted to the inter-participant variability in brain structures. Consequently, dipoles fitted to those scalp maps will be more accurately localized in the 3D space of brain model used (Onton & Makeig, 2006; Oostendorp & van Oosterom, 1989).

Of course, ICA is not necessary for calculating the 3D location of an electrical source. The source of electrical activations can be predicted based on the scalp maps from non-ICAd electrode data as well. This is similar to identifying the electrodes which are optimally positioned to record the signal of interest based on the structural properties of the brain for each participant, and then clustering these electrodes. To our knowledge, this sort of spatial optimization is not frequently used: Typically, an electrode is identified *a priori* and the analyses are conducted on the data from that electrode.

To sum up, we believe i) clustering equivalent neurological processes based on 3D location of their generator is more accurate than on electrode location, and ii) ICA leads to more accurate estimates of the location of these generators then using electrode data. Further, factoring in of other established qualities of an EEG signal (e.g. power at a given frequency, or polarity) in the clustering function on top of the expected dipole location makes the process of identifying the ICs of interest more sensitive and specific. Because all of these parameters (theta power, 3D coordinates of the dipole, latency of the event, etc.) are quantitative, all of this can be achieved without compromising the objectivity and general validity (i.e. replicability) of the methods (c.f. selection of ICs by eye).

The ERN is a very robust EEG signal with well-established polarity, latency, scalp map, and theta oscillatory properties. However, we still had a number of novel difficulties in clustering the equivalent ICs. These difficulties are outlined below, along with a description of semi-automated approach we took to tackle them, and a description of how we checked the validity of our solutions to these problems.

Clustering the Present Data

The aim of the cluster analysis was to blindly allocate equivalent ICs into separate clusters based on a set of parameters (e.g. power and inter-trial coherence of theta oscillations at the time of error key-press, see chapter 5 for a description of how these parameters were extracted). At the end of the cluster analysis, the ICs which are furthest away from each other in the space of these parameters are placed in different clusters. Similarly, ICs which are closest to each other in the space of these parameters are placed in the same cluster. As such, the within-cluster variance of ICs for these parameters would be minimal; and between-clusters variance would be maximal. For example, if we were to set the dipole location of the ICs as the only parameter of interest, the clustering algorithm would allocate all the ICs such that all the ICs in a given cluster would be maximally close to each other in physical space within the brain, and maximally far from ICs in other clusters.

We knew from inspection of individual participant data and also from previous literature that ErICs i) had dipole locations close to ACC (Carter et al., 1998), ii) were associated with high theta power and ITC peaking just before the error key-press (Trujillo & Allen, 2007) and iii) that these effects would be most pronounced during corrected errors (Hewig, Coles, Trippe, Hecht, & Miltner, 2011). Thus we knew that dipole location and theta oscillations at the time of key-press should be two parameters to be weighted heavily in the clustering process, and we should use the corrected error epochs in the clustering process.

Problems The first problem was to decide how many clusters to generate. The k-means clustering algorithm built into EEGLAB requires the number of clusters to be set before the cluster analysis starts. The second problem was that even though we knew what parameters to use to identify the error components, we didn't know how much weight each parameter should be given when constructing the clusters. These two problems were exacerbated by the lack of any set guidelines, or systematic approaches to the problem in the literature. Thus, deciding how many clusters to generate or how much weight each parameter should be given in the clustering process was difficult and partially uninformed.

Solution Our approach to this problem was in many ways trial and error. We set up a loop in MATLAB to plot the outcomes of all possible combinations of weights (ranging from 0 to 8), and cluster numbers from 5 to 20, and plot the results. We found that 8 clusters was enough to reveal a cluster which contained all the error related components. The optimal combination of parameter weights was as follows: Strong weights (8 out of 8) for theta power at the time of key-press (-50 to 200ms after the key-press) and dipole location; 0 weight for the ERP (0 out of 8) and medium weight on ITC (3 out of 8). We acknowledge that this is a rather exploratory approach to setting the parameters, but in the absence of any precedents which would guide us, we reasoned that this was the most objective and replicable approach we could use.

Checks on the Validity of Solution One specific observation boosted our confidence in this approach. Even though the algorithm used to cluster the independent components was blind to the ERP amplitude, the resulting cluster with ErICs showed a clear and reliable ERN and a Pe for corrected but not uncorrected errors (see the figure 4.3, for a summary of the error related cluster properties). This adds empirical support for the idea that ERN is associated with frontal theta oscillations, and more importantly for our purposes, shows that the independent components in this cluster are indeed those that show ERP changes in response to errors, as would be predicted from the literature. From this point on we refer to the cluster which contained the ErICs as the error related cluster.

The average scalp maps, dipole locations, ERPs, event-related spectral perturbations (ERSP) and inter-trial coherence (ITC) values of the clusters are shown in figures 2.2, 2.3, 2.4, 2.6 and 2.5 respectively. These figures suggest that the ErICs were successfully clustered together in cluster 6. Please note that cluster 1 in these figures are not shown. This is because the cluster 1 is the parent cluster, which contains all of the ICs included in the cluster analysis.

Error related cluster was composed of 15 ICs from 11 of the 12 participants, which meant one participant contributed no ICs to the error related cluster. On more detailed inspection of EEG data of this participant, we found that none of the ICs yielded by ICA showed error related changes such as the ERN or theta oscillation changes. To further understand if this was because of bad ICA resolution, we looked at the changes in the data from fronto-medial electrodes. We found no error related changes recorded by these electrodes. Thus, we were convinced that the absence of an error related IC for this participant was not due to the failure of the clustering

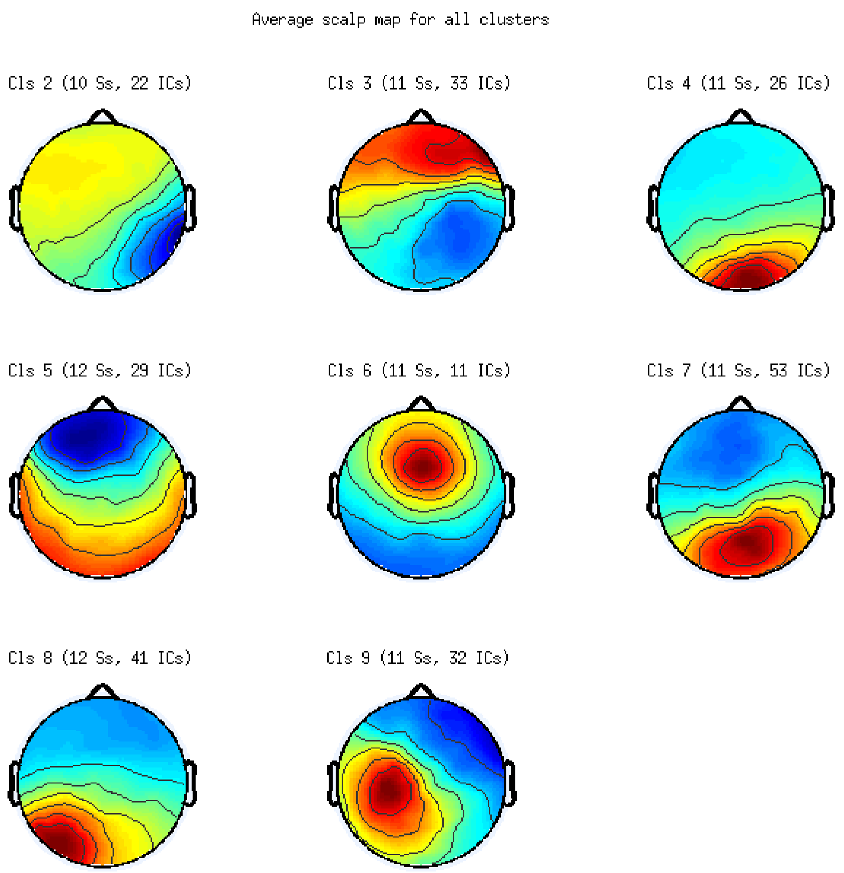


Figure 2.2: Average scalp maps of the clusters revealed by the cluster analysis

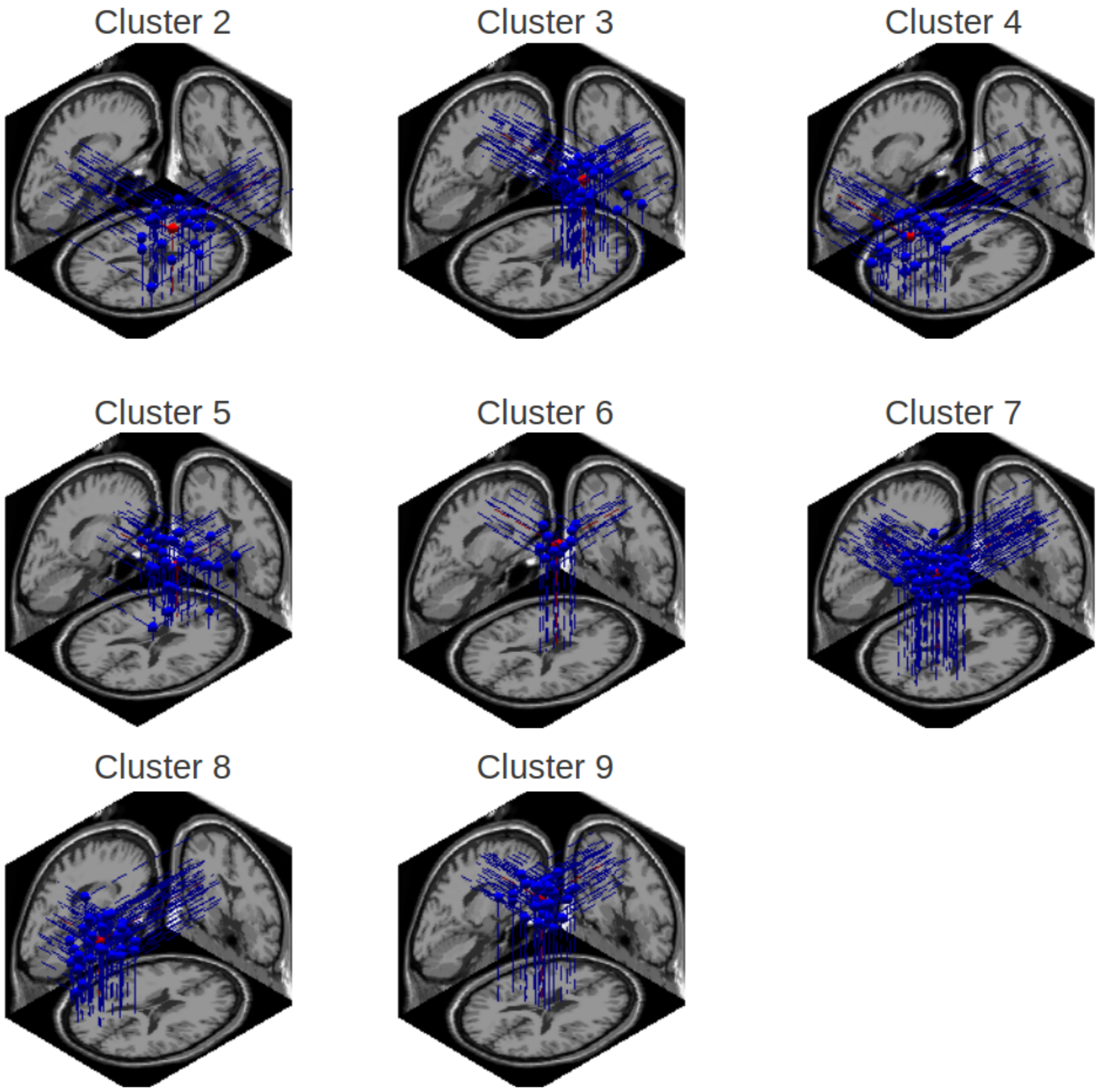


Figure 2.3: Location of IC dipoles within each one of the clusters revealed by the cluster analysis

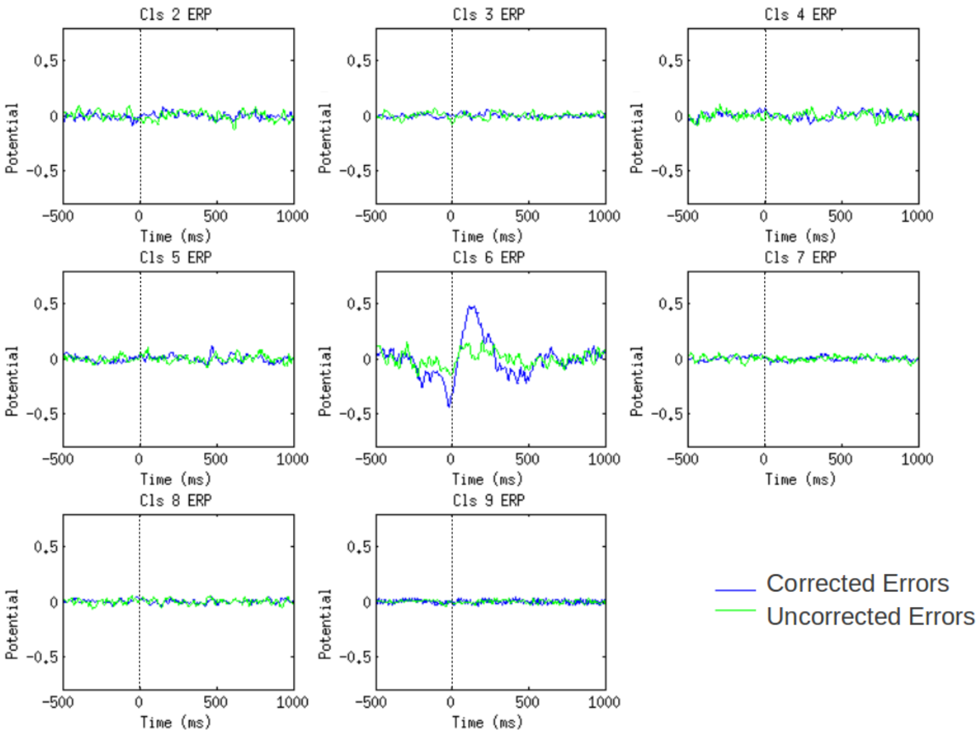


Figure 2.4: Average ERPs (in micro-volts) of the clusters revealed by the cluster analysis. Time zero, marked by the dashed lines shows the time of error key-press.

algorithm to identify an ErIC, or bad ICA decomposition, but rather due to lack of error related changes in the electrode data.

Four participants contributed 2 ICs to the error related cluster. For these participants, dipole locations, scalp maps, ERPs and theta power and ITC were inspected and the IC which was clearly not related to errors was removed from the error related cluster, leaving the cluster with 11 ErICs from 11 participants.

These observations suggested that clustering algorithm and the parameters/weights used were sensitive enough to detect at least one ErIC from all participants who showed error related EEG activity, but also specific enough to exclude all ICs from the participant who did not show any error related EEG activity as predicted from the literature.

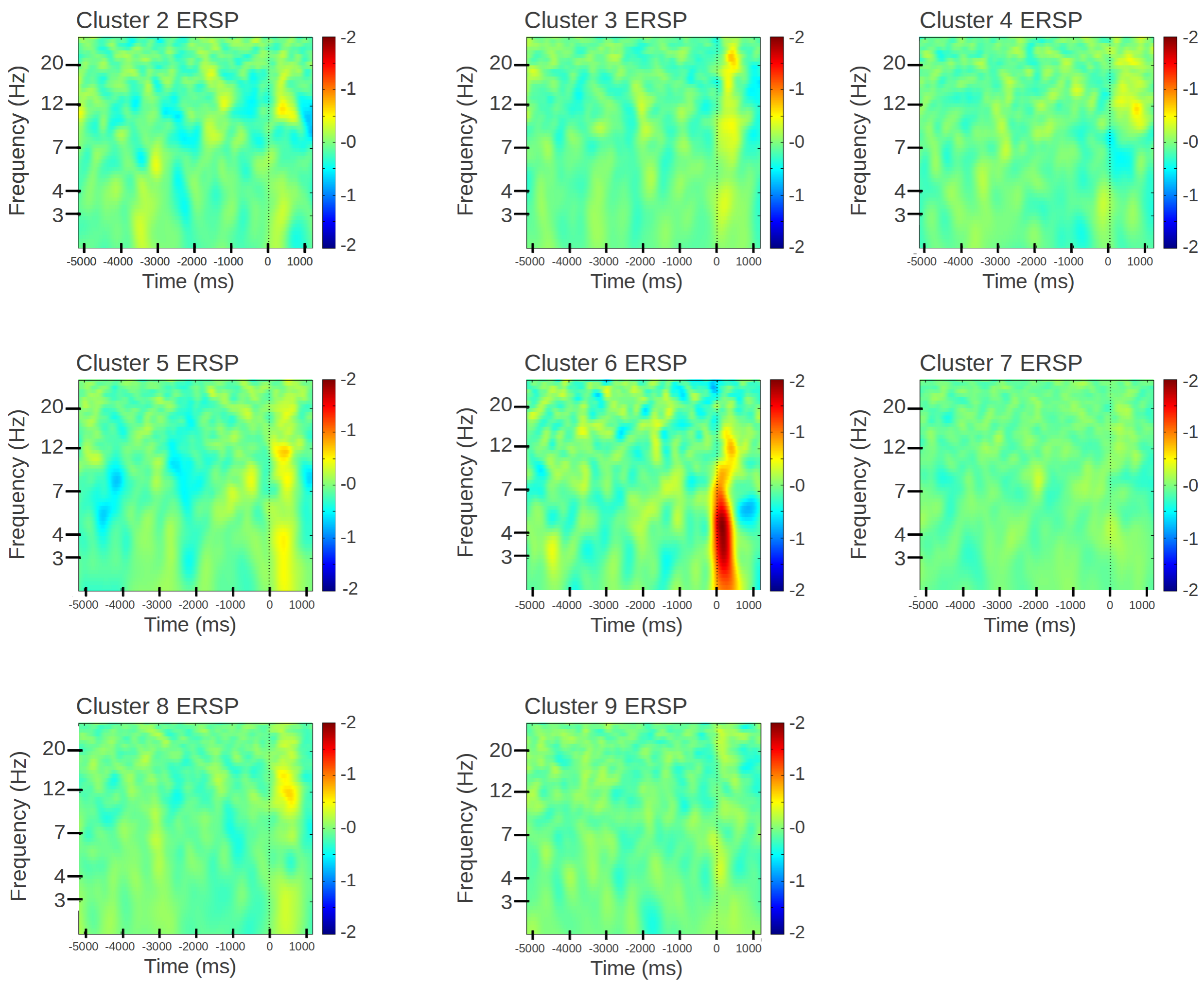


Figure 2.5: Average ERSP values of the clusters revealed by the cluster analysis. Magnitude of ERSP (in decibels), is represented by the colour scale.

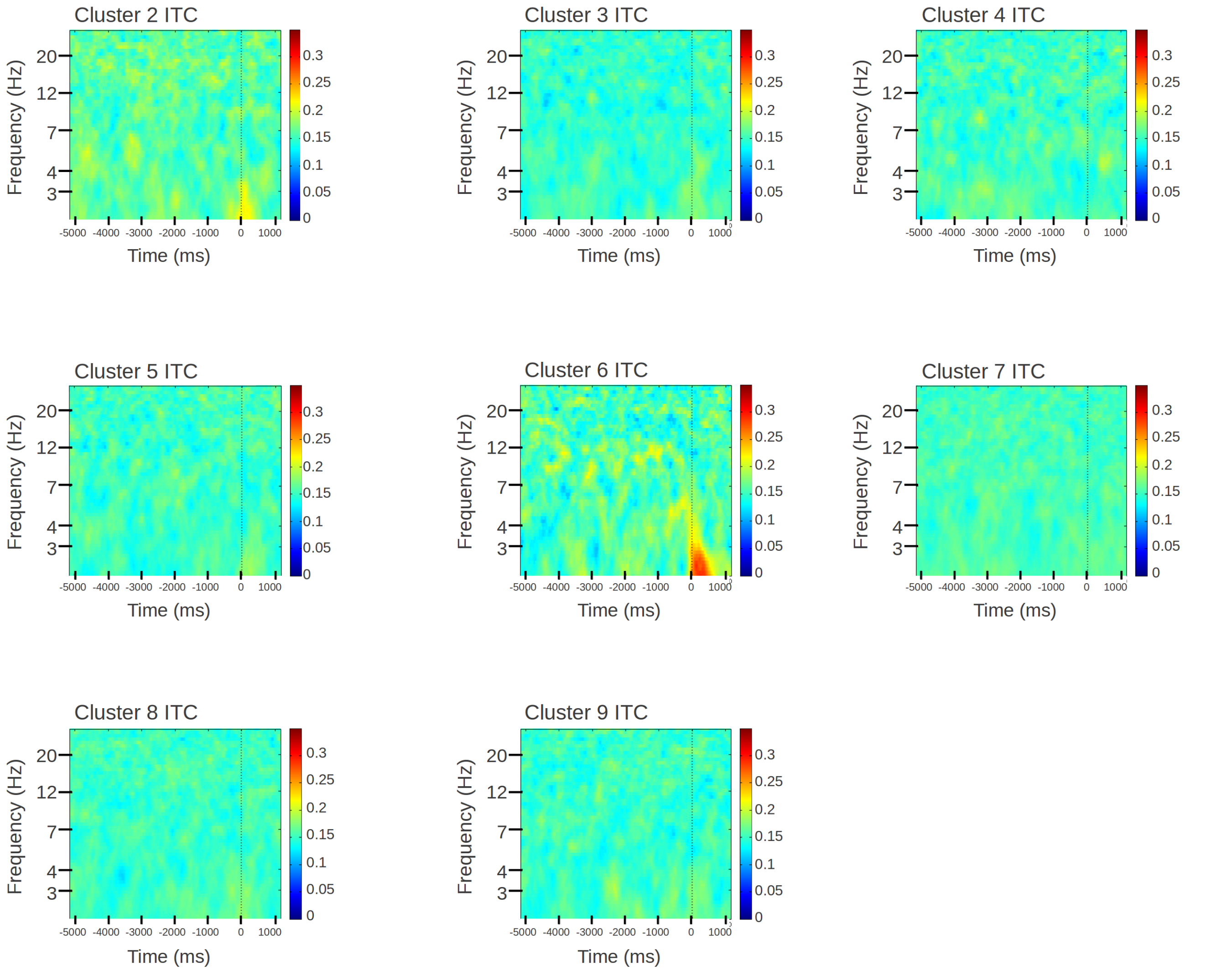


Figure 2.6: Average ITC values for each cluster revealed by the cluster analysis, represented by the colour scale.

Chapter 3

Experiment 1

3.1 Introduction

Performance monitoring is an important skill for making the necessary adjustments to ongoing behaviour in response to changes in environment or long term goals. Insight into performance monitoring can be gained through the study of errors. To avoid confusion, we use the term cognitive control systems to refer to the mechanisms responsible for keeping ongoing behaviour in line with long term goals and changes in environment (c.f. 'outer loop' (G. Logan & Crump, 2009, 2010); or the 'supervisory attentional system' (Norman & Shallice, 1986)). We use the term performance monitoring to refer to the mechanisms which signal the need for increased cognitive control. The need for a performance monitoring system arises particularly for the control of highly practised tasks involving precise coordination of very quick movements.

3.1.1 Typing

Touch-typing has a number of benefits as an experimental paradigm for the study of psychological processes, which were already recognized many decades ago (Lashley, 1951; Wells, 1916). Firstly, typing behaviour is a natural action, which has become an integral part of people's professional and social lives. Because of this, the number of hours of practice an ordinary person acquires over several years in typing is close that elite athletes or musicians acquire in their fields (Ericsson &

Krampe, 1993). It shares common aspects with other skilled actions such as driving or musical performance. A key similarity is the chunking together of multiple action units through practice. Once sufficient practice has been undertaken, the amount of cognitive effort required to execute these actions becomes minimal (Norman & Shallice, 1986). However, this does not mean that the importance of monitoring performance in these actions becomes less important, particularly in tasks such as driving. Performance in these actions is subject to intervention from cognitive control systems and can be adjusted as external stimuli change or internal goals are updated (G. Logan & Crump, 2011).

Usually, one problem with bringing such ecologically valid tasks into the psychology lab is the lack of control the experimenter has over it, and the difficulty in quantitatively evaluating the accuracy of performance. For example, there are lots of different ways of making tea, or getting a car from where it is to a certain point. In typing however, there is only one way of typing any word correctly. The correct letters should be typed in the correct order. Because any violation of this rule constitutes an error, the distinction between errors of performance and accurate performance is clear. This is an important advantage when studying performance monitoring, because most of the performance monitoring literature is focused on error and post-error performance, and it would be difficult to interpret our results in this context without a clear *a priori* description of errors. The fact that typing involves hundreds of finger presses every minute (Rosenbaum, 1991) ensures that a lot of error instances can be observed within a relatively short amount of time (cf. discrete trial choice reaction time (CRT) tasks de Bruijn et al., 2003; Holroyd, Dien, & Coles, 1998; Miltner, Braun, & Coles, 1997). We believe typing is a suitable method for studying natural behaviour in the psychology laboratory, because what the participant needs to do is tightly controlled by the presented text, and any deviance from the text constitutes an error.

3.1.2 Performance Monitoring

As mentioned in the introduction chapter, performance monitoring mechanisms have been studied using experimental paradigms including different versions of flankers (Ullsperger & von Cramon, 2006; van Veen & Carter, 2002), go/no-go (Scheffers et al., 1996) and Stroop tasks (Vidal et al., 2003, 2000), which involve discrete responses. In almost all of these experiments, the focus is on the errors the participants make, and their results show that error and pre-error responses are faster, and post-error responses are slower than the correct responses (Laming, 1979; Rabbitt, 1966b, 1968). Increased speed in error and pre-error trials can be explained by speed-accuracy trade-off: The faster you are, the more likely you are to make mistake (Wickelgren, 1977). Similarly, one may argue that post-error slowing serves to bring performance speed back to a level where accuracy is almost certain (i.e. to compromise speed to achieve accuracy). However, the response times in trials immediately after the error are slower than response times associated with the highest likelihood of accuracy (Rabbitt & Rodgers, 1977). This shows that post-error performance is slowed down more than necessary to achieve optimal performance. Such over-compensatory effects may reflect the engagement of performance monitoring processes.

Skilled typing involves hundreds of finger key-presses every minute. This results in a large number of error instances, even with relatively low error rates. In addition to engaging performance monitoring mechanisms, errors inform us about what parameters are important for accurate performance. A direct comparison of errors and correct responses in typing may reveal important differences not only after an error, but also before it. Just as post-error effects can inform us about performance monitoring mechanisms, pre-error changes can inform us about the interaction between other cognitive processes and motor output processes. For example, we may find that the amount of pre-error variance in a given measure (e.g. force) predicts accuracy better than that in others (e.g. speed). To our knowledge, pre-error changes in skilled actions have not received much interest in the performance monitoring literature.

How skilled typists monitor their performance is particularly interesting because they can type tens of letters in a matter of seconds with relatively low error rates. Because most finger movements in typing are initiated before the previous ones in the sequence are completed (Flanders & Soechting, 1992), typists usually have no conscious insight about which letter they are typing at any given moment. However, they can signal their errors almost instantly (G. Logan, 1982; Rabbitt, 1978). It is curious how typists can judge the accuracy of each finger movement when they don't know where their fingers are at a given time.

According to Logan and Crump (2011), typing behaviour is controlled in a hierarchical way. There are two components (or loops) involved in copy-typing: The outer loop is involved in converting the visually presented text into linguistic units (i.e. words), and then passing these units to the inner loop. The inner loop translates these units into individual letters and eventually to key-presses. Further, the outer loop relies on feedback from the ultimate outcome of the typing action, the output on the screen. The inner loop on the other hand, relies on somato-sensory feedback from the fingers, and is not affected by the output of the screen (G. Logan & Crump, 2010). Even though this model provides us with predictions about what type of feedback might be used in the correction of errors in typing, it is not designed to explain how that feedback is used or represented in monitoring performance.

3.1.3 Detection of Errors and Response Times

There are multiple hypotheses accounting for the detection of errors. However, most of these hypotheses were developed to explain how errors are detected when there are only two response alternatives (see Alexander and Brown (2010) for a review of these models). Compared to the amount of research that has gone into error detection mechanisms in non-continuous and non-skilled actions, error detection mechanisms in continuous and skilled actions such as typing have received little interest.

Most studies using single trial, choice reaction time (CRT) tasks show that people detect their errors post-response: Post-error slowing is a well established result (Rabbitt & Rodgers, 1977). After participants make a mistake, their response time in the following trial is very slow. The error trial itself however, is faster than average (Rabbitt, 1966b). This result suggests a feed-back mechanism for error detection, where errors are detected after the error action is initiated. This interpretation is strengthened by neuro-imaging studies. Many EEG studies have now confirmed that there are time-locked changes in the activity of frontal areas of the brain very shortly after the onset of the error response (i.e. ERN, Falkenstein et al., 1991; Gehring et al., 1993).

However, a number of studies using skilled, continuous tasks have found that people slow down *before* making an error and execute error responses with less force compared to correct responses (Herrojo-Ruiz et al., 2009; Rabbitt, 1978; Shaffer, 1975). Shaffer (1975) showed that in typing, error keys were slower than correct keys, and some keys preceding the error (i.e. pre-error keys) were also slower than keys preceding correct keys. Rabbitt (1978) showed that error key-presses in typing were pressed down with less force than correct key-presses. Herrojo-Ruiz et al. (2009) have shown that in piano playing, key-presses up to 3 keys before errors are slower than those preceding correct keys. Ruiz and colleagues further reported the onset of the ERN to precede that of error response.

Herrojo-Ruiz et al. (2009) claim that the pre-error slowing and pre-error ERN (or pre-ERN) can be explained by early error detection based on feed forward models (Wolpert & Miall, 1996). One hypothetical function of feed forward models is to predict the sensory outcome of a motor command before that motor command is executed by the effector muscles. According to Herrojo-Ruiz et al. (2009), the error can be detected ahead of time because 1) skilled actions such as piano-playing involve preparation of multiple responses ahead of the time of execution, and 2) these responses can be compared to the correct actions, and thus any mismatch can be detected *before* a response is initiated. This interpretation of pre-error slowing suggests a feed-forward mechanism for error detection. We refer to this account linking pre-error slowing to error detection as the *early error detection* hypothesis.

3.1.4 Pre-Error Slowing - Cause or Effect

Based on the contrasting observations from CRT tasks (where error and pre-error actions are fast) and skilled, continuous tasks (where the error and pre-error actions are slowed), there are two competing hypotheses explaining pre-error performance. Speed accuracy trade-off account predicts that speeding up should increase likelihood of error commission, decreasing accuracy. One testable hypothesis based on this is that participants' errors and pre-error actions should be faster compared to correct actions. On the other hand, early error detection account predicts that because errors can be detected before they are executed, pre-error responses should be slowed down compared to correct actions.

Another explanation for pre-error slowing also acknowledged by Herrojo-Ruiz et al. (2009) is pre-error performance breakdown. This alternative explanation suggests that it is not necessarily the error detection that leads to slowing down of the key-presses, but that the relationship can also be the other way around: Performance starts to degrade, as indexed by slowing down, loss of rhythm etc., and this foreshadows error commission. One advantage of typing over piano playing on this issue is that typing errors are naturally signalled by the backspace key. Thus, using backspace we can separate errors which were detected from those which were not, isolating the effect of error detection. Further, if pre-error slowing in typing is due to error detection, we should be able to see it only before detected errors.

3.1.5 Aims and Predictions

The primary aim of the current study was to test two predictions about pre-error performance. Namely, 1) The speed-accuracy trade-off hypothesis' prediction that errors should be preceded by faster key-presses than those before correct key-presses, and 2) Early error detection hypothesis' prediction that key-presses just before the error should be slowed down compared to those before error key-presses.

In addition, we wanted to explore the relationship between error detection and variance in pre-error performance. Based on subjective reports of participants and also on previous research suggesting that changes in pre-error mental states might be related to error commission and detection (Cavanagh et al., 2009; H. Eichele et al., 2010; T. Eichele et al., 2008), we investigated the variability in the pre-error key-presses. Many participants commented that when they are 'in the zone' of typing, they did not think too much about the task, and their fingers typed the words smoothly with little cognitive effort. We reasoned that when the participants are not 'in the zone' of typing their key-presses should consist of more abrupt pauses or delays compared to when they are in the zone. This would be expressed as higher variability in typing performance. Thus, a third prediction we tested was that the errors preceded by smaller variance in typing speed will be more likely to be detected and corrected than those preceded by higher variance.

3.2 Methods

Twenty one participants who had been trained as touch-typists (formally through a course, or informally using typing software) took part in the study. Participation in study was voluntary and took 15 to 20 minutes in total. Two of 21 participants were left-handed, and 4 were males, with an average age of 35 (range 20 - 63). The average typing speed for participants was 66.95 words per minute (SD = 15.16) with an average accuracy of 94.55% (SD = 6.39). Seven of the participants were students and the rest worked for the University of Sheffield in different departments and libraries. Informed consent was obtained prior to the start of the experiment, in line with University of Sheffield ethics regulations.

3.2.1 Design

Participants typed 20 sentences presented on a computer screen: Five of the sentences were headlines from BBC website, 5 were quotes obtained from the website `www.quotesoftheday.com`, 5 were sentences from the journal *Science*, and 5 were song lyrics obtained from various websites on the internet. Each sentence was made up of 13.8 (SD = 2.33) words on average, longest sentence composed of 17 words and the shortest of 10 words. Each word on average was composed of 4.7 letters. Flesch-Kincaid Reading Ease score (Kincaid, Fishburne, Rogers, & Chissom, 1975)

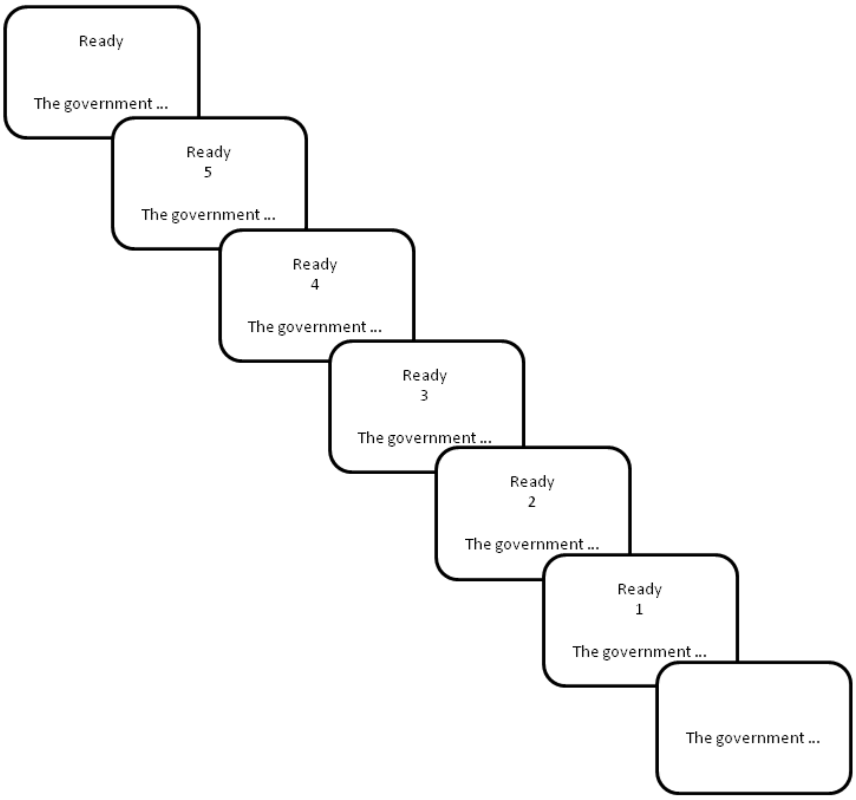


Figure 3.1: An example trial from the experiment. Participants were asked to read the sentence while the word 'Ready' was on the screen in order to avoid reading and typing at the same time.

of the sentences on average was 60.1, with a grade level of 8.3. The order of presentation of these sentences was randomized for each participant. See Figure 3.1 for an example trial.

3.2.2 Procedure

After signing the informed consent form, the participants were given verbal instructions by the experimenter. The participants were asked to type as fast and accurately as possible, correcting any mistakes made. They were told they would not get any visual feedback (i.e. what they typed would not be shown on the screen), and the view of their hands and the keyboard would be blocked. The instructions were also presented on the computer screen before the first trial.

Each trial began with the presentation of a sentence and a simultaneous visual countdown (Figure 3.1). The word 'Ready' stayed on the screen for 2 seconds before the first number appeared and each number stayed on the screen for one second.

Participants were asked not to start typing the sentences before the countdown was over, but that they should use this time to read the sentence once before starting to type instead. The speed of sentence presentation was self paced: Once they finished typing, the participants had to press the '+' button on the number pad to begin the next trial. All stimuli were created and present using MATLAB©psychtoolbox.

3.2.3 Analysis

Error Slowing As described in the methodology chapter, we recorded the time and identity of each letter typed online and calculated the IKIs offline. We controlled for confounding factors of letter position and word length while calculating the slowing associated with error, pre- and post-error slowing. As a reminder, error slowing was calculated for each error key-press IKI by subtracting the average IKI of correct keys matched for letter position and word length:

$$\text{Error Slowing} = \text{IKI}_{\ Error\ Keypress} \text{ - } \text{IKI}_{Avg.\ Matched\ Correct\ Key-presses}$$

The same procedure was applied to calculate the slowing associated with the letters preceding (pre-error slowing) and following (post-error slowing) the error letters. Average error slowing was calculated separately for each participant in order to control for individual differences in typing speed. We use the following abbreviations to refer to pre- and post-error key-presses. For the error key, we use 'E'; for the key that immediately follows the error key, we use 'E+1'; for the key that immediately precedes the error key we use E-1, and so on, such that the 'E-6' refers to the key-presses executed 6 keys before the error, and 'E+3' refers to that executed 3 key-presses after the error key. Pre- and post-error key-presses were only considered when they were in the same word as the error (G. Logan & Crump, 2011).

Following Rabbitt (1978) we hypothesized that corrected errors and uncorrected errors would show different patterns of IKI changes compared to correct key-presses. For this reason, we looked at the effect of corrected and uncorrected errors on the IKI separately.

Error Types

We excluded omission and insertion errors (i.e. simultaneous key-presses) from our analysis. We defined errors which were constituted by a missing letter from an otherwise correct word as omissions. The reason we excluded them is that 1) a key not pressed doesn't have an IKI, and 2) we had no way of confirming whether these were genuine errors or caused by a key-press not being recorded by the keyboard. Insertions were those errors with an IKI of less than the 5^{th} percentile and adjacent to the preceding or the following key-press. The reason we excluded these insertion errors is that these errors are likely to be caused by a single finger movement leading to the pressing of two keys on the keyboard almost simultaneously. Because we use IKIs to represent one single finger movement, we excluded any errors which were caused by a finger pressing two keys simultaneously. All the remaining error types were collapsed together.

Treatment of Outliers and Missing Data

We took the following precautions to minimize the effect of outliers and the small number of observations on the statistical analysis.

First, for each participant, we identified the IKI corresponding to the 99^{th} percentile, and excluded all IKIs slower than that. Second, we used non-parametric statistical methods. Non-parametric methods are preferred to parametric counterparts when the data are not normally distributed. In our case, especially for early pre-error key-presses such as E-5 and E-6, the number of observations was smaller than others, simply because not all error key presses were preceded by 6 key-presses. The smaller the number of observations in a given distribution, the less likely will its shape be normal. Thus, instead of using t-tests which assume data are symmetric around the mean, we used Wilcoxon's signed ranks test. Because this non-parametric analogue of the t-test relies on the ranks of the data, it is much less affected by potential outliers than the t-test (Howell, 2002).

Calculation of confidence intervals around the average error slowing values was based on bootstrapping technique (Howell, 2002), and used in making inferences about p values (using 10000 re-samples). For example, when the 95% confidence intervals excluded zero the p-value is reported to be < 0.05, when 99% confidence intervals exclude zero p-value is reported to be < 0.01, and so on.

Error Detection

The analyses described so far are based on error slowing, which is a measure of the difference between error performance and accurate performance. However, as mentioned in the Aims and Predictions section, we were also interested in differences between corrected and uncorrected errors (i.e. the effect of error *detection*).

Classification of Corrected vs. Uncorrected Errors We labelled those errors which were corrected by a backspace as 'corrected' errors, and those errors which the participant continued typing without pressing the backspace as 'uncorrected' errors. We are interested in inferring the cognitive process of error detection from the behavioural measure of error correction but recognise that this is not straightforward. This is why we use the terms 'corrected' and 'uncorrected' (as opposed to 'detected' and 'undetected'). It is not possible to be certain that all uncorrected errors were actually not detected. It is probable that the participants became aware of at least some of these errors, but were too late/lazy to act on them. Another possibility is that the participant actually pressed the key, but not strong enough for the keyboard to register it. Objectively, this would be recorded as an omission error, but from the participant's perspective, it would not be an error at all. We tried to work around this problem by excluding omission errors from our analysis, but we are nevertheless unable to ultimately confirm whether each 'uncorrected' error is indeed 'undetected'. We therefore interpret with caution this difference between corrected and uncorrected errors, acknowledging that error correction is, at best, an imperfect approximation of error detection.

By looking at the differences between uncorrected and corrected errors immediately before the error (e.g. E-2 and E-1), we tested predictions about the effects of error detection on performance. Similarly, by looking at earlier keys (e.g. E-6 to E-3), we tested predictions about the effect of pre-error performance on error detection.

Error Correction Rate and Pre-Error Slowing If error detection leads to pre-error slowing, then the errors associated with higher pre-error slowing should be more likely to be corrected. To test this prediction, for each participant, we divided the errors into two groups based on the amount of pre-error slowing. The errors preceded by low pre-error slowing (average error slowing at E-2 and E-1 faster than the 33^{rd} percentile of all E-2 and E-1 key-presses) were separated from those with high pre-error slowing (average error slowing at E-2 and E-1 slower than the 66^{th} percentile of all E-2 and E-1 key-presses). For each group, we calculated the proportion of corrected errors, resulting in two rates of error correction for each participant: The rate of error correction for errors with high pre-error slowing and that for cases with low pre-error slowing.

Error Correction Rate and Pre-Error Variance We measured the variability in the pre-error key-presses to see if it was related to error correction. We used E-6 to E-3 key-presses for this analysis because key-presses immediately before the error (E-2 & E-1) are shown to be affected by error commission by previous research (Herrojo-Ruiz et al., 2009; Shaffer, 1975). We used *error slowing* rather than raw IKIs because this measure controls for the variance caused by word length and letter position on IKIs. As the measure of variance, we used the 68.2% inter-percentile range as a non-parametric analogue of standard deviation (see the next paragraph). For each error key-press, we extracted the pre-error slowing values (E-6 to E-3) and calculated their inter-percentile range. Then we compared the amount of variance before corrected errors to that in uncorrected errors. Further, we collapsed the corrected and uncorrected errors, and separated them into the ones preceded with the highest amount of pre-error variance and those preceded with the lowest amount of pre-error variance, the 33^{rd} and 66^{th} percentiles as cut-offs. This allowed us to check whether errors preceded by high pre-error variance were less likely to be corrected than those preceded by low pre-error variance.

We used the 68.2% inter-percentile range rather than the standard deviation, because 1) the IKI and error slowing distributions did not have symmetric distributions, and 2) one standard deviation around the average accounts for the middle 68.2% of the distribution, and so does the 68.2% inter-percentile range.

Note on statistical power for pre-error variance analysis

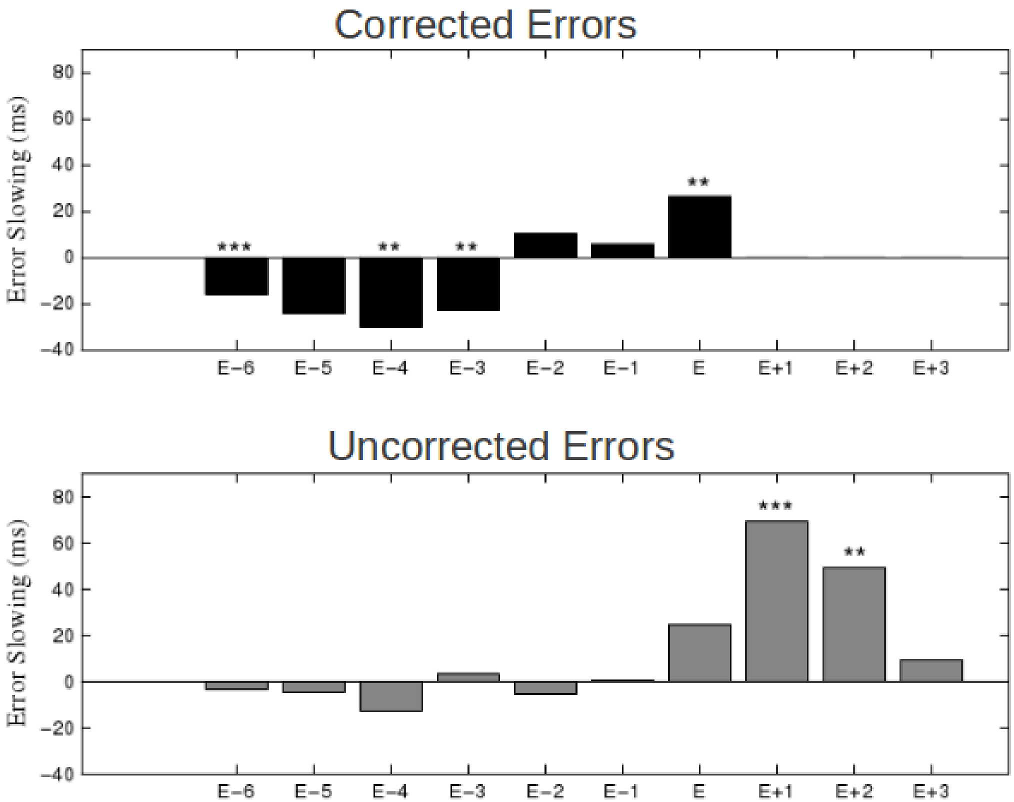
We found that the sample sizes for the pre-variance analyses were quite small for some participants (On average each participant contributed 4 instances of errors to this analysis). A closer look at the data showed that this was because 1) earlier pre-error key-presses (e.g. E-6) were less numerous than later ones (e.g. E-3), 2) Unless there are at least 2 data points (i.e. unless there are at least 4 keys before an error), a variance score can't be calculated. In other words, in order for a pre-error variance score to be calculated for an error, it must be preceded by at least 4 correct key-presses (E-3 & E-4).

3.3 Results

3.3.1 Post-Error Results

The mean error slowing times across all participants are outlined in Figure 3.2. Post-error slowing is only presented for uncorrected errors. We found that 2 key-presses following an error were significantly slowed down compared to matched correct key-presses. Mean post-error slowing associated with E+1 and E+2 keys were 69.57ms (SD = 88.53, $p < 0.001$), and 49.50ms (SD = 88.00, $p < 0.01$) respectively.

We do not report post-error slowing for corrected errors because the number of correct key-presses typed by participants was very low. Of the 19 participants, did not press any correct post-error keys before the backspace. On average each of the 15 participants contributed 2 (SD = 1.25) post-corrected-error key-presses.



** - 95% Confidence intervals exclude 0*
** - 99% Confidence intervals exclude 0*
*** - 99.9% Confidence intervals exclude 0*
Note. Bootstrapped confidence intervals for the mean are based on 10000 re-samples

Figure 3.2: The bar chart shows how much slower, in milliseconds, the key-presses in incorrectly typed words were compared to matched key-presses in the correctly typed words.

3.3.2 Error Slowing

Corrected Errors

We found that corrected error key-presses were typed with a delay of 26.39ms (SD = 39.96) compared to correct key-presses ($p < 0.01$).

Uncorrected Errors

Uncorrected errors were no more slower than matched correct key-presses. Average error slowing = 24.55ms (SD = 97.76, $p > 0.05$)

3.3.3 Pre-Error Slowing

Corrected Errors

Key-presses preceding the corrected errors were found to be typed *faster* than those preceding matched correct keys. This pre-error speeding was particularly reliable in E-6 (M = 16.08ms, SD = 20.31, $p < 0.001$), E-4 (M = 30.00ms, SD = 39.27, $p < 0.01$) and E-3 letters (M = 23.24ms, SD = 39.99, $p < 0.01$). As can be seen in Figure 3.2, there seems to be a sudden disappearance of this pre-error speeding 2 key presses before the error. To see if this was a reliable change in speed, we compared the average error slowing across all participants in E-6 to E-3, to that in E-2 and E-1. A Wilcoxon's signed ranks test showed that pre-error slowing in E-6 to E-3 was much smaller (M = -24.03ms, SD = 32.52) than that in E-2 and E-1 (M = 8.14ms, SD = 36.75), Z = -3.33, $p < 0.001$. We refer to this relative 32.16ms slowing down 2 keys before the corrected errors as 'local pre-error slowing'.

Uncorrected Errors

We found no pre-error slowing effects for uncorrected errors (all p's > 0.05). As Figure 3.2 suggests, there was no indication of local pre-error slowing before uncorrected errors.

3.3.4 Error Detection and Pre-Error Slowing

Pre-error Slowing in E-2 and E-1

When we compared the average pre-slowing in E-2 and E-1 in corrected errors (M = 8.14ms, SD = 36.75) to that in uncorrected errors (M = 1.54ms, SD = 57.02) using a Wilcoxon's signed ranks test, we failed to find a reliable difference (Z = -0.48, p = 0.63).

Pre-Error Slowing in E-6 to E-3

A Wilcoxon's signed ranks test showed that the amount of pre-error slowing in E-6 to E-3 in corrected errors (M = -24.03ms, SD = 32.52), was not significantly different from that in uncorrected errors (M = 6.10ms, SD = 73.31), Z = -1.02, p = 0.309.

Local pre-error slowing

We calculated local pre-error slowing for each error committed by each participant by subtracting the immediate error slowing (E-2 & E-1) from early error slowing (E-6 to E-3). Then we compared the average of local pre-error slowing for each participant in corrected errors to those in uncorrected errors. Wilcoxon's signed ranks tests showed that average local pre-error slowing was not statistically different than zero at the participant level neither for corrected errors (M = 11.55ms, SD = 29.46, Z = -1.50, p = 0.133) nor for uncorrected errors (M = -29.25ms, SD = 83.71, Z = -1.25, p = 0.212). Further, the amount of local pre-error slowing in corrected words was not significantly different from that in uncorrected errors (Z = -1.73, p = 0.084).

3.3.5 Pre-error slowing and Error Detection

As described in the *Analysis* section, we calculated the error correction rate for errors preceded by high and low pre-error slowing for each participant. We found that how much a participant slowed down before an error did not affect the probability of error correction. This was true for early (E-6 to E-3), immediate (E-2 & E-1) and local pre-error slowing measures (all p's > 0.05).

3.3.6 Error Detection and Pre-Error Variance

We compared the amount of variance before corrected errors to that of uncorrected errors as described in the *Analysis* section. A Wilcoxon's signed ranks test showed that the amount of early pre-error variance (based on E-6 to E-3) was reliably smaller in corrected errors (M = 78.80ms, SD = 37.22) than that before uncorrected errors (M = 132.66ms, SD = 48.54), Z = -2.62, p = 0.009.

3.3.7 Error Detection Probability and Pre-Error Variance

A Wilcoxon's signed ranks test showed the errors preceded by high variance in performance were much less likely to be corrected (M = 49.02%, SD = 41.59) compared to those preceded by low variance (M = 71.93%, SD = 46.94), Z = -2.06, p = 0.038.

3.4 Discussion

3.4.1 Summary of Results

The primary aim of our study was to investigate the pre-error performance in a skilled and continuous action, touch-typing. We found that errors of typing do *not* cause pre-error slowing in typing. In fact, our participants' corrected errors were foreshadowed by faster key-presses than average. To our knowledge, pre-error speeding in a continuous task is a novel finding in the skilled actions literature.

We found that corrected, but not uncorrected error key-presses were slowed down compared to matched correct key-presses. This supports the results of Shaffer (1975) who showed that error key-presses were slower than correct key-presses, and those of Rabbitt (1978) who showed that signalled errors were executed with less force (or 'pulled up') compared to non-signalled errors. We also found that errors which were not corrected were followed by slowed key-presses (post-error slowing). This replicates the results of G. Logan and Crump (2010) who showed that typing is slowed down after undetected errors, and extend it from typing of single words to copy-typing in a continuous manner in the absence of visual feedback.

3.4.2 Effect of Error Detection on Pre-Error Performance

Our results show that changes in pre-error speed before corrected errors are not different than that before uncorrected errors. This suggests that the early error detection interpretation of pre-error slowing (Herrojo-Ruiz et al., 2009) does not generalize to typing. We believe that there are a number of reasons for this contrast between Herrojo-Ruiz et al. (2009)'s and our behavioural results.

First, there is a lack of a natural behavioural error signalling response in piano playing (c.f. the backspace in typing). This makes it difficult to assess error detection objectively and precludes disentangling the effects of *error commission* from *error detection*.

Second is a very important distinction between piano playing and typing. Piano-playing performance is constrained to an extra dimension compared to typing: In typing, any finger movement is correct as long as it leads to the typing of the letter that needs to be typed: Pressing the key 'a' when the letter 'a' needs to be typed will be accurate, irrespective of the speed of typing, or the force applied when typing it. Whether this key-press is too fast, or too strong will not compromise its accuracy. In piano playing however, the accuracy of performance depends on the timing and the force of the key-press as well as the note that needs to be played. Any violation of what is dictated by the score, particularly in terms of the timing/speed of key-presses will constitute an error. Thus, unlike typing, speed of performance in piano playing performance is strictly constrained by external rules.

Herrojo-Ruiz et al. (2009)'s participants were asked to play at a speed of 8 notes per minute [or an inter-onset-interval (IOI) of 125ms]. However, these pianists slowed their IOIs from an average of 121ms in correct key-presses to an average of 190ms up to 3 key-presses before making a mistake. This amounts to a delay of more than 50% for each one of the 3 keys pressed before the error. Given this amount of change from what is dictated by the score, it is possible that the pianists considered at least one of these 3 slow pre-error keys as incorrect, in at least some of error instances. This would not only cause error slowing (disguised as pre-error slowing), but also a 'pre-ERN': Olvet and Hajcak (2009) show that 6 to 8 pure instances of errors are enough to obtain a statistically significant ERN when compared to a set of correct responses. Thus, it is plausible that an incorrectly slow execution of a correct note preceding an incorrect note yields a partial, but nevertheless reliable ERN.

Because of these reasons, we believe that typing is a more sensitive task than piano playing for studying changes in performance speed before the error. Accordingly, we reject the interpretation that pre-error slowing in skilled actions is caused by error detection. Our finding of a local, rather than an overall pre-error slowing across participants might seem to support the pre-error slowing - error detection relationship, particularly because it was only observed before corrected errors. However further analyses of our data revealed that, 1) local pre-error slowing was not observed at the participant level, 2) the amount of local pre-error slowing before corrected errors was no different from that before uncorrected errors, and 3) it did not increase the likelihood of error correction. Accordingly, we believe that any error commission or detection effects on performance surfaces during or after, but not before the initiation of the error action. Further, two reliable effects in our data point to changes in early pre-error performance as a factor affecting error detection probability, rather than the other way around. These effects are described next.

3.4.3 Effect of Pre-Error Performance on Error Detection

We found that the only factor which reliably affected the rate of error correction was the variability of the typing speed. We found that errors committed while the participants were 'in the zone' (as one participant described it), were more than 25% more likely to be corrected than otherwise. This result can be interpreted in at least two possible ways.

One interpretation is to assert that there are two inherently different types of errors. One type is caused by going very fast (pre-error speeding), and consistently so (low pre-error variance), and these errors are more likely to be corrected. The other type is caused by a performance breakdown as indexed by increased variability and loss of rhythm (high pre-error variance), and is harder correct. According to this interpretation, one kind of error is in its essence more likely to be detected by the typist than other.

Another interpretation is that errors don't systematically vary in how detectable they are, but rather are more likely to be detected when the participant is in an alert mental state. In other words, when the participant is alert and her 'performance monitoring system is engaged' 1) her typing will be fast and smooth, and 2) she will be more likely to correct her mistakes. According to this interpretation, typists are more likely to detect their mistakes if their performance monitoring systems are engaged.

Our study was not originally designed to differentiate between these two accounts, and more research involving neural measures is clearly needed to provide a satisfactory answer. However, our analyses provide enough information to at least offer some speculation on the matter.

First, taking a step back, we are not aware of any theoretical or empirical work which predicts that the errors caused by performance breakdown should be inherently harder to detect than those caused by an increase in speed. Speed-accuracy trade-off relationship predicts high speed will lead to more errors, but makes no predictions about the detectability of these errors.

Second, we find that error commission, irrespective of explicit error detection and correction, causes slowing (e.g. post-error slowing after uncorrected errors). This suggests that most errors are registered at some level. Importantly, the timing of this slowing is affected by error detection. In corrected errors (which were preceded by much more consistent typing) slowing starts soon after the error initiation (as hinted by slowed error key-press execution). This is what one would expect if the typist was in an alert state with a highly engaged performance monitoring system. In uncorrected errors (which were preceded by highly variable and un-rhythmic key-presses) however, slowing down starts only after the error has been committed (post-error slowing). Again, this is what one would expect if the typist was not in an alert state and their performance monitoring mechanisms were not engaged.

Based on these two points, we speculate that when the typists are typing more smoothly, they are more likely to detect their mistakes, and, that the rhythmicity and consistency in finger presses may be indices of mental processes involved in motor production as well as performance monitoring.

However, another explanation we can't exclude is that errors which are committed when typing simpler (i.e. simple to type) or more familiar words are more likely to be detected. This would also lead to the observation of more consistent and smooth typing performance preceding the corrected errors compared to uncorrected errors. We tried to work around this problem by controlling for word length and letter positions, but none of these two measures can directly control for the familiarity effect. Since the words familiar to each participant will be most likely different to those familiar to other participants, this was a difficult problem to tackle. This problem is exacerbated by the small number of pre-error key-presses obtained. A more direct way to assess the hypothesis that the error detection is driven by the alertness/attentional state of the participant is to use neuro-imaging (chapter 5).

3.4.4 Hierarchical Control in Typing: Implications of The Current study

As mentioned in the introduction, G. Logan and Crump (2011)'s hierarchical control of typing involves two loops, which depend on different kinds of feedback. The inner loop is sensitive to the kinaesthetic/proprioceptive feedback from the fingers, and the outer loop is sensitive to the final product of the typing behaviour as it appears on the screen. Further, the inner loop is informationally encapsulated, such that the outer loop does not know how the inner loop gets the job (typing of individual letters) done.

Within this framework, our participants had no feedback at the outer-loop level, because the screen provided no feedback on the participants' typing and the workings of the inner loop are not accessible to the outer loop. However, in many instances, our participants made mistakes in the middle of words, pressed the backspace, and started from the right position in the word. If the outer loop doesn't know what the inner loop is doing, and has no feedback other than the screen, then it cannot instruct the inner loop to stop typing, press the backspace, and continue from where the error was initially committed.

Our results suggest that the outer loop does have access to different channels of feedback, and these channels are weighted differently under different circumstances. For example, during everyday typing where the typist can see the output of his performance, the outer loop relies almost exclusively on the unambiguous visual feedback. G. Logan and Crump (2010) have shown that this is the case even when typists are told that the visual feedback they get from the monitor can be misleading during the experiment. Even under these circumstances, typists judged the accuracy of their own performance based not on sensory feedback from their fingers, but on the potentially 'untrustworthy' monitor. Observations from our study show however, that when the only available sources of feedback are the relatively noisy (c.f. the visual feedback from the screen) sensory information from the fingers, the outer loop *will* exploit these sources. We believe that with this addition to Logan and Crump's hierarchical model of typing, our results are compatible with it.

3.4.5 Conclusions

In conclusion, we show using a simple and ecologically valid task that error detection doesn't lead to pre-error slowing in skilled actions. On the other hand, error instances which were corrected and post error actions were indeed slowed, revealing the effect of performance monitoring processes *after* the error action is initiated. We show that in a continuous task, pre-error variance in performance is predictive of error detection and can be a behavioural marker of performance monitoring processes.

Chapter 4

Experiment 2: Time Domain Analysis of Errors

4.1 Introduction

The aim of this chapter is to report EEG as well as behavioural findings from our typing experiment. The importance of performance monitoring and the measures used to study it have been already discussed in the introduction chapter of the thesis. To avoid duplication, only the specific topics investigated in this experiment will be covered in the introduction section of this chapter.

One of our primary aims in this experiment was to replicate and extend the error related EEG components such as the ERN and Pe, and then to establish their relationship to error awareness and behavioural changes (i.e. error and post-error slowing) in the skilled and continuous action of typing. In addition we aimed to replicate the behavioural findings we reported in the previous chapter with more data (we used 100 sentences in this study rather than 20 in the previous one).

4.1.1 Error Awareness and Error ERPs

In the introduction chapter, we stated that the relationship between the ERN and error awareness was not well established with many studies failing to find a significant correlation between ERN amplitude and the participant's awareness of error commission. Wessel (2012) recently reviewed the articles which studied the relationship between ERN and the participants awareness of errors. Wessel's review reports a total of 15 publications, 8 of which report a significant correlation (i.e. $p < 0.05$), 7 of which report no correlations (i.e. $p > 0.05$).

All but one of these studies used discrete trial tasks. Hewig et al. (2011) on the other hand, asked their participants to enter 5 digit numbers using a custom build number pad in every trial. Participants were given no visual feedback, were asked how confident they were about the accuracy of their response (correct, unsure, incorrect) at the end of the trial.

Hewig et al. (2011) found that among objectively error trials, the average ERN amplitude was the largest for subjectively error trials, intermediate for subjectively unsure trials, and smallest for subjectively correct trials (all contrasts statistically reliable). Further, when compared to the correct trial ERPs, ERN was only significant during subjectively incorrect, and subjectively unsure trials, but not during error trials which were judged to be correct. Hewig et al. (2011)'s results suggest that in a more complex action such as entering numbers on a number pad, error awareness might be related to ERN amplitude. To our knowledge, this is the only performance monitoring study which investigated the relationship between EEG measures and error awareness in a continuous response (i.e. digit entry) task.

4.1.2 Independent Components and ERPs

Before applying ICA to our data, we did not have clear predictions about whether one IC would be representing both ERN and Pe or separate ICs would be revealed for ERN and Pe by ICA. This is because ERN has been observed after almost all errors, but Pe is only observed after errors which are detected (see section 1.2.3 for a review). This contrast suggests that ERN and Pe might be manifestations of temporally independent mechanisms and thus separated into distinct ICs by ICA.

On the other hand, previous performance monitoring studies which decomposed EEG data using ICA have consistently reported that the IC associated with error responses exhibited Pe as well as ERN (e.g., Gentsch et al., 2009; Hoffmann &

Falkenstein, 2010; Murphy et al., 2012; Roger et al., 2010; Wessel et al., 2012). Thus, the current empirical evidence from ICA studies of ERN suggest that ERN and Pe are not temporally independent processes, and are manifestations of a single process.

4.1.3 Aims of the current study

The aims of the current study can be summarized in three points. First, we aimed to replicate and extend the EEG findings of ERN and Pe to skilled and continuous actions using typing as an experimental paradigm. Herrojo-Ruiz et al. (2009)'s findings suggest that in skilled actions such as piano-playing, the onset of ERN might precede the onset of error response due to internal forward models. Hewig et al. (2011) haven't found any such effects during digital entry of numbers on a number pad.

Second, we aimed to assess the relationship between error awareness and ERN and Pe. Typing is well-suited for this purpose because typing errors that reach the awareness of the typist will be signalled by an implicit and well integrated component of the typing actions repository, the pressing of backspace. Because back-spacing is such a well learned sub-action in typing, there is no need to ask participants to stop or report whenever they thought they made an error.

Third, we aimed to study the correlations between the amount of error slowing or post-error slowing and the EEG measures such as the ERN and Pe. Error slowing and post-error slowing in particular have been used as indices of error detection during typing in the past (G. Logan & Crump, 2010; Rabbitt, 1978; Shaffer, 1975).

4.2 Methods

4.2.1 Participants

In total, 21 participants volunteered for the study, 9 of whom were males. The mean age was 29 years (SD = 7.64). Participants were invited to participate by email, and included students, librarians and secretaries associated with different departments of the University of Sheffield. Informed consent was obtained prior to the start of the experiment, in line with University of Sheffield ethics regulations.

4.2.2 Design and Procedure

The design and procedure for the experiment are described in section 2.4, so will not be repeated here.

4.2.3 Behavioural Task

Participants were required to type the first 100 sentences from the first part of the book Cumulative Record (Skinner, 1959). The order of presentation of the sentences was randomized for all participants. Sentences were presented on a computer screen preceded by the word 'Ready' which stayed on the screen for 2 seconds (see figure 4.1). After typing the sentence, participants had to press the right arrow key on keyboard to start the next sentence, and were told that they could use this as an opportunity to rest between the sentences if they needed to. Participants had no visual feedback as their hands were covered and the output of their typing did not appear on the screen. MATLAB Psychtoolbox was used to present the sentences and record the key-presses.

4.2.4 Data Analysis

Behavioural Analysis

Analysis of behavioural effects such as error slowing followed the exact same steps as those in the experiment reported in chapter 3.

Selection of Matched Correct Key-presses for EEG Analysis

As mentioned in the methodology chapter, we were interested in the EEG contrast between correct and error key-presses. To this end, we created an algorithm to identify correct key-presses matched to error key-presses, based on IKI and letter. For example, if the first error key-press was an *e* with an IKI of 150ms, the

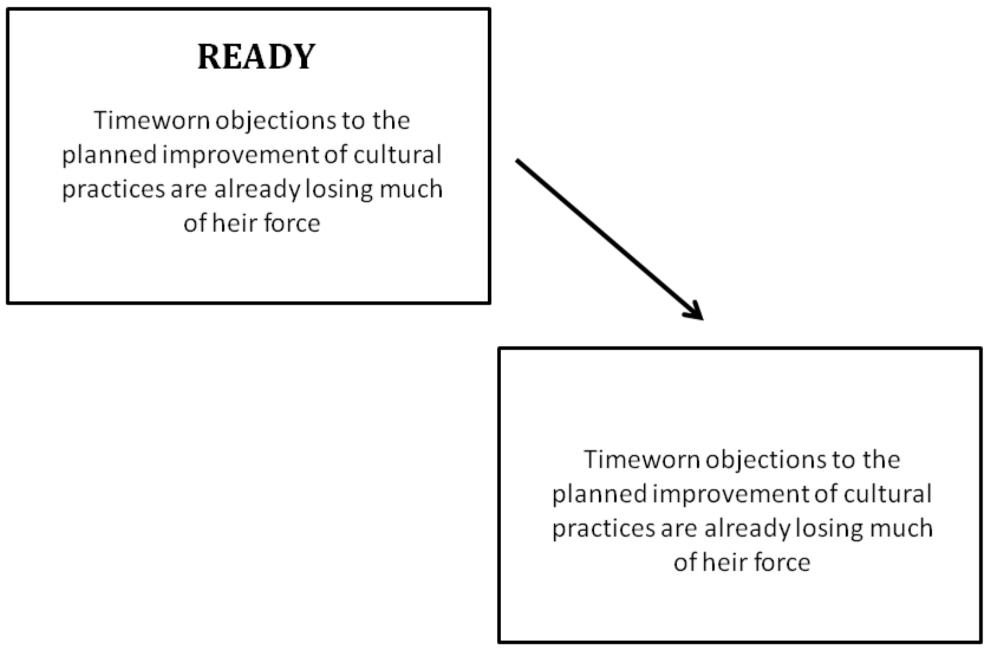


Figure 4.1: Figure showing the presentation of the sentences. The participants were asked to start typing after the 'Ready' sign disappeared.

algorithm would identify all the correctly typed *e*'s and identify the ones with an IKI of 150. A staircase procedure was applied to find the closest IKI with a step size of ± 0.5ms. Note that this is a different matching procedure than that used for behavioural analysis, where key-presses were matched for letter position and word length.

EEG Data Acquisition and Analysis

On-line Data Acquisition EEG data were collected from 128 channels at a sampling rate of 2048Hz using a Biosemi Actiview system (`http://www.biosemi.com/`) referenced to the electrode A1. Data were analysed by custom Matlab scripts built on the open source EEGLAB toolbox (Delorme & Makeig, 2004, `http://sccn.ucsd.edu/eeglab/`)

Off-line Data Processing and Analysis The EEG data were down-sampled from 2048Hz to 256Hz using Biosemi BDF Decimator software (`http://www.biosemi`

.com/). The rest of EEG data analysis followed the below steps: 1) The data were digitally filtered to remove frequencies above 60Hz and below 1Hz using a finite impulse response filter, as implemented in the pop_eegfilt function in EEGLAB. 2) Key-press events in the EEG record were labelled as *corrected error, uncorrected error, correct matched to corrected error* and *correct matched to uncorrected error.* 3) Continuous EEG data were cleaned of artefacts and noisy channels by visual inspection. Artefacts were defined as sudden and substantial ($>$ 3 standard deviations) changes of amplitude in multiple electrodes at the same time. Blinks, swallows and head movements are among the typical generators of such artefacts. If these substantial changes appeared only in one electrode with no similar activations in spatially adjacent electrodes, the electrode was identified as a potentially noisy channel. If similarly large and isolated changes appeared in the same electrode multiple times, that electrode was identified as a noisy electrode and removed. After removing noisy channels, 109 (SD = 13.44) channels were retained on average for each participant. 4) The data were re-referenced to the average electrode. 5) Continuous data were submitted to extended infomax ICA (Lee, Girolami, & Sejnowski, 1999) using runica function (Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997) of the EEGLAB toolbox. The average number of data points decomposed for each participant was 533985 (34 min, 46 sec). 6) Independent component source locations were estimated by creating an equivalent current dipole model for each component using dipfit function from EEGLAB. This function estimates dipole location by applying inverse source modelling methods to a standard boundary element head model (Oostendorp & van Oosterom, 1989). 7) The ICs whose dipoles had a residual variance of more than %20, or were outside the brain were removed. Any remaining components that reflected muscle activity, electrocardiogram, or eye movements, on the basis of their dipole location, spectra and scalp maps were considered artefacts and excluded from further analysis. In total, 251 ICs were included in the analysis, each participant contributing 21 (SD = 6) ICs on average. 8) EEG data were than separated into epochs of corrected errors, uncorrected errors, correct matched to corrected errors and correct matched to uncorrected errors. Epoch length was

8 seconds long starting from 6 seconds before to 2 seconds after the key-press. 9) Finally, after pre-computing their ERPs, power spectra, ERSPs and ITC values at the time of key-press, all remaining ICs from all the participants were submitted to a cluster analysis, using the k-means clustering function in EEGLAB (as described in section 2.4.4). Extraction of ERSP and ITC are explained in more detail in the methods section of chapter 5. Briefly, extraction of oscillations from EEG data was accomplished using wavelet analysis using Morlet wavelets (Herrmann, Grigutsch, & Busch, 2005) as implemented by the newtimef function in EEGLAB toolbox. Baseline period used for ERSP and ITC calculations was from 6000ms to 200ms before the key-press.

Single Trial ERP Analysis We were interested in 2 different kinds of analysis involving EEG measures. First, the ERP analysis - to replicate and extend ERN and Pe results to the skilled action of typing and document their temporal relationship to the errkey-press. Second, the correlation between these ERPs with behavioural measures (error and post-error slowing). The second analysis requires quantification of measures such as the ERN amplitude and latency for each trial. This in turn requires that a certain time-window is set, the negative peak within which is designated as the ERN amplitude. Then the latency of this negative peak can be used as the ERN latency.

As can be seen in figure 4.8, the ERPs of error key-presses started to diverge from those of correct key-presses before the error key-press reaching a negative peak at 19.53ms before the key-press in our data. The onset of this divergence (i.e. the earliest time point at which error key-press ERPs were significantly different than correct key-press ERPs) was found to be 156.25ms before the error key-press and lasted until 19.53ms after. Thus, for the single trial correlational analyses, the latency and the amplitude of the negative peak within this time-window was used as the latency and amplitude of ERN within trials (as Falkenstein et al. 1991 did). Following the same steps, the time window for quantifying Pe peak amplitude and latency was found to be 70.31ms to 210.94ms after the key-press. Positive peak and its latency within this time-window was used as the Pe amplitude and latency for single-trial correlational analyses.

The selection of the critical time period for between trial correlations analysis was driven by our own data rather than previous literature simply because we are not aware of any EEG studies which investigated trial-by-trial correlation analysis using a skilled and continuous task. Because the latency of ERN in continuous actions such as typing is not established, and the closest findings from piano playing suggest the ERN latency might actually precede the actual error key-press (Herrojo-Ruiz et al., 2009), we refrained from defining an *ad hoc* time window. Rather, we used visual inspection of cluster analysis results to guide our definition of ERN and Pe. A similar approach has been used by Debener et al. (2005, p. 11732) to determine the boundaries of the critical time window in a discrete trial task.

In the paragraph below we present previous approaches and measures used to define the critical period and the peak amplitude in an effort to provide a comparison of our methods to previous literature.

Falkenstein et al. (1991), publishing the first instance of ERN, reported the time window used to reveal it to be between 50ms before to 150ms after the response. Gentsch et al. (2009) similarly used the difference between the most negative peak in the time window of 0 to 100ms after the key-press and the most positive peak between 100 to 0ms before the key-press. Debener et al. (2005) used the difference between negative peak within the 15 to 85ms window after the key-press and the mean of the two positive peaks surrounding it (i.e. the positive peaks in -80ms to 0ms and 85ms to 240ms). These methods are representative of the larger literature, and show that in discrete response tasks like the flankers task, the ERN is expected to peak within the first 100ms of response onset.

It is important to clarify here that the behavioural measure used in these correlational analysis is error slowing, and not error IKI. Error slowing is not an absolute measure but a relative measure, indicating how much slower an error IKI is compared to a matched correct IKI (matched to letter position and word length). This ensures that the behavioural measure used here is already baseline corrected.

4.3 Results

One of the 21 participants was excluded from the experiment due to bad typing data: This participant found it very difficult to type in the absence of visual feedback, frequently shifted her hands from the home position without realizing it. Thus her key-press data were not usable neither for behavioural analysis nor labelling the responses in EEG. Testing of another participant had to be cancelled before the EEG recording started because the Biosemi System failed to start. During the testing session of one participant, EEG amplifier stopped working and no EEG data could be collected. But her key-press data were saved and used for the behavioural experiment. Of the remaining 18 participants 5 were excluded due to a large number of noisy EEG channels and too many artefacts, and one was removed because he had less than 10 epochs of undetected errors after artefact removal (Olvet and Hajcak (2009) show that a minimum of 6 to 8 instances of error epochs are necessary to reveal a statistically reliable ERN). This left the number of participants in the EEG analysis at 12 and the number of participants in the behavioural analysis at 19. Average baseline typing speed and accuracy for the 19 participants were 70.37 words per minute (SD = 13.32), and 94.21% (SD = 4.85), respectively.

4.3.1 Behavioural Results

The average IKI (of all key-presses excluding the backspace) across all participants was 155.81ms (SD = 19.43). The effect of errors on the IKIs of error, pre- and post-error key-presses are shown in figure 4.2.

Error Slowing

Corrected Errors Corrected error key-presses were found to be 37.22ms (SD = 13.72, $p < 0.001$) slower than matched correct key-presses.

Uncorrected Errors Uncorrected error key-presses were also found to be significantly slower than matched correct key-presses (M= 22.31ms, SD= 16.98, $p < 0.001$).

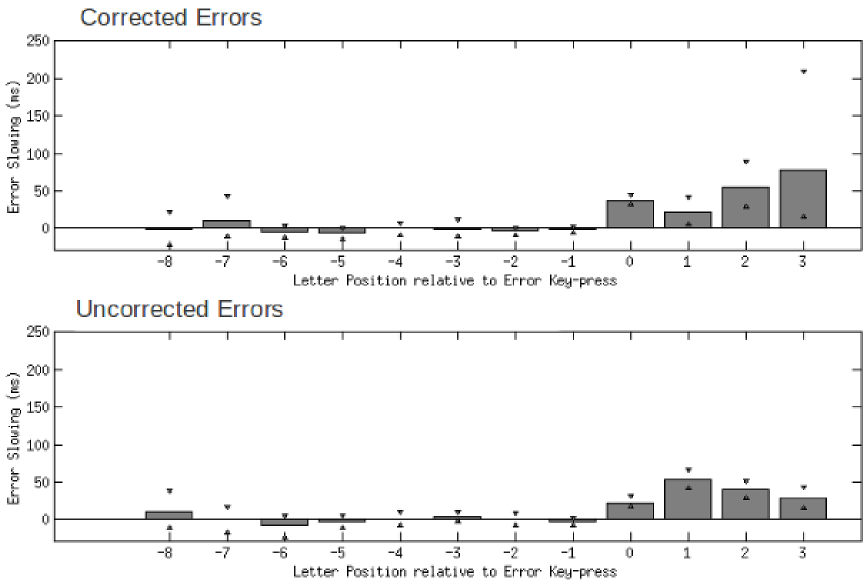


Figure 4.2: Amount of Error Slowing associated with error, pre- and post-error key-presses for corrected (top) and uncorrected (bottom) error key-presses. On the x-axis, the numbers show letter position relative to the error, 0 being the error key-press, -1 being the key-presses immediately before the error, and 1 being the key-press immediately following the error. Arrow heads show 95% confidence intervals

Post-Error Slowing

Corrected Errors The amount of post-error slowing associated with corrected errors were 20.99ms (SD= 39.09), 54.13ms (SD= 58.22) and 78.03ms (SD = 110.90), for E+1, E+2 and E+3, respectively. Post-error slowing values were all statistically reliable (all p's < 0.01)

Uncorrected Errors Post-error key-presses following the uncorrected error key-presses were found to be significantly slowed down compared to matched correct key-presses. The average post-error slowing was 52.72ms (SD= 29.46), 40.62ms (SD= 26.57), and 27.52ms (SD= 33.11), for E+1, E+2, and E+3, respectively (all p's < 0.001).

Pre-Error Slowing

We found no reliable slowing before neither corrected, nor uncorrected errors.

4.3.2 EEG Results

We visually summarized the error cluster in a series of figures below. Figure 4.3 shows the average scalp map, dipole locations, ERSP, ERP, and ITC measures of the error related cluster, and figures 4.4, 4.6 and 4.5 show the scalp maps, ERSPs and ERPs for ErICs contributed to the error related cluster by individual participants.

Note on Unequal Number of Epochs

We calculated the number of observations in the four different kinds of key-presses that was included in the EEG analysis. We found that participants contributed 51.25 (SD = 26.44) corrected error, 79.08 (SD = 45.26) correct matched to corrected error, 25.58 (SD = 13.28), uncorrected error and 41.00 (SD = 21.89) correct matched to uncorrected key-press epochs to the EEG analyses. The number of observations in the 3 comparisons made (i.e. corrected error vs. matched correct, uncorrected error vs. matched correct, corrected vs. uncorrected errors) were found to be statistically different from each other by t-tests (all p's < 0.05), which caused concerns.

The disparities in the numbers of epochs were caused by our study being not strictly experimental in design. Natural typing involves many more correct key-presses than error key-presses, and typists can detect most of their mistakes (G. Logan, 1982; Rabbitt, 1978). Participants made 12053 key-presses on average, 390 of which were errors on average. This amounts to about 30 correct key-presses for every error key-press. After the matching procedure we had less than 2 matched correct key-presses for each error key-presses, which were not significantly different in terms of IKI from the error for any of the participants (all p's > 0.05). Even though we acknowledge that the number of error observations could be matched to correct key-presses, there was not much we could do about the disparity between the corrected and uncorrected error key presses: People simply corrected most of their errors, and we actually asked the participants to do this in order to ensure that when a participant misses an error (i.e. doesn't correct it), it is not because they were too lazy or careless, but because they really were not aware of making an error. Thus, other than excluding certain corrected errors to bring their numbers down to match the number of uncorrected errors, we can not conceive of a way of equalizing the number of corrected and uncorrected error observations.

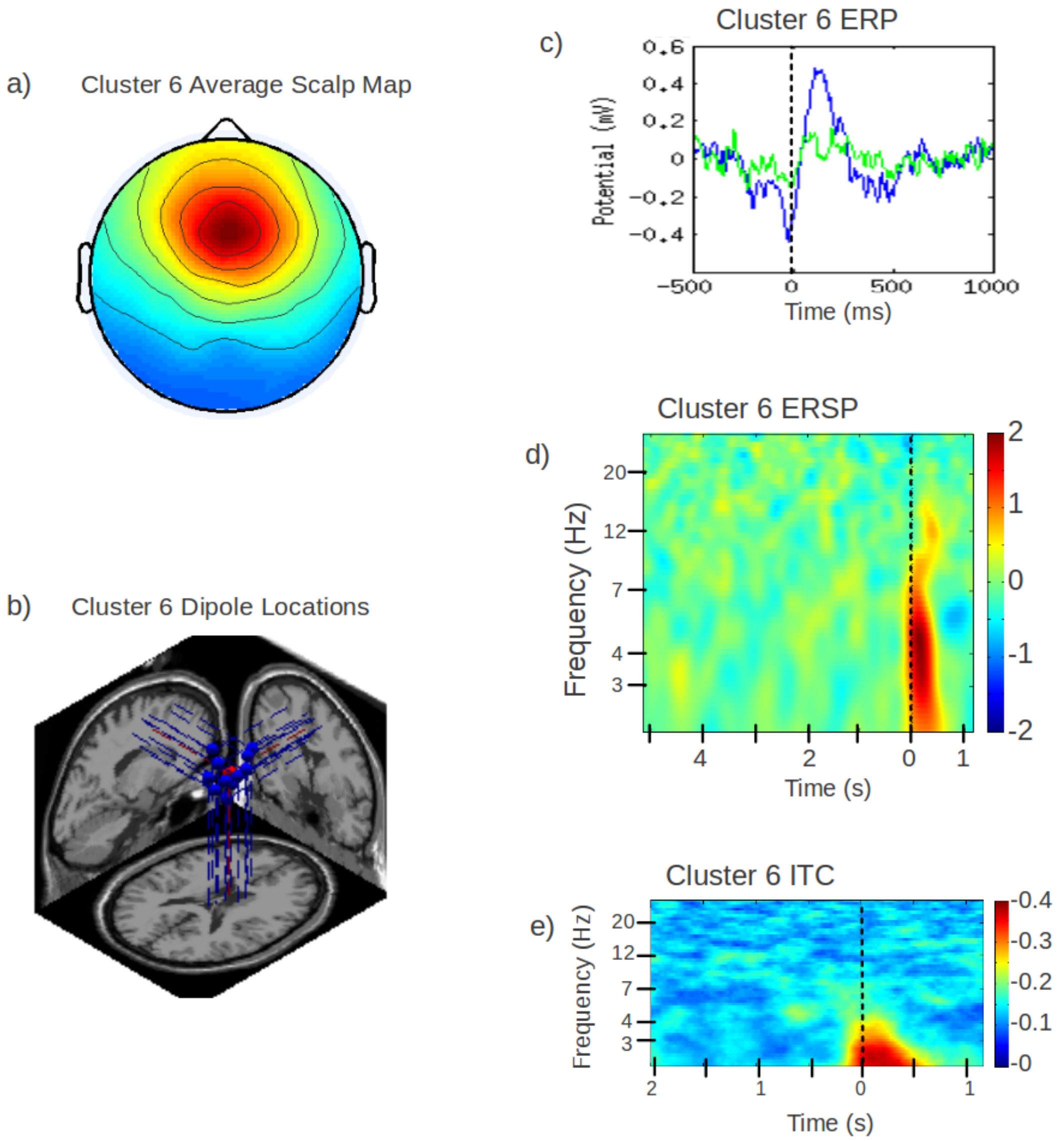


Figure 4.3: The average scalp map (a) and dipole locations (b, average dipole in red) of the error related cluster. A comparison of ERP activations (c) of the ICs during corrected error key-presses (blue line) to matched correct key-presses (green line). Average ERSP (d) and ITC (e) values of the ErICs in the error related cluster at the time of corrected error key-presses. For d) and e) the magnitude of ERSP and ITC are represented by the colour scale.

Corrected Error Key-presses vs. Matched Correct Key-press

On average, the ICs in the error related cluster showed significantly different response locked activations in response to corrected error key-presses compared to matched correct key-presses. As shown in figure 4.8, the ERN for corrected errors

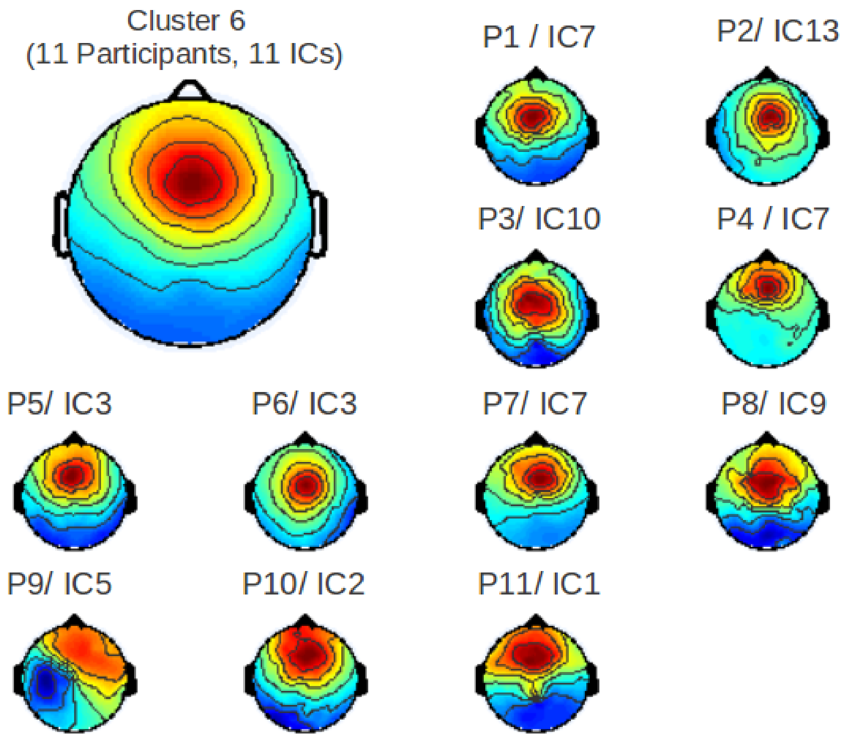


Figure 4.4: The average scalp maps of all the ErICs in the error related cluster (large, upper left), and the scalp maps for the ErICs contributed by each participant

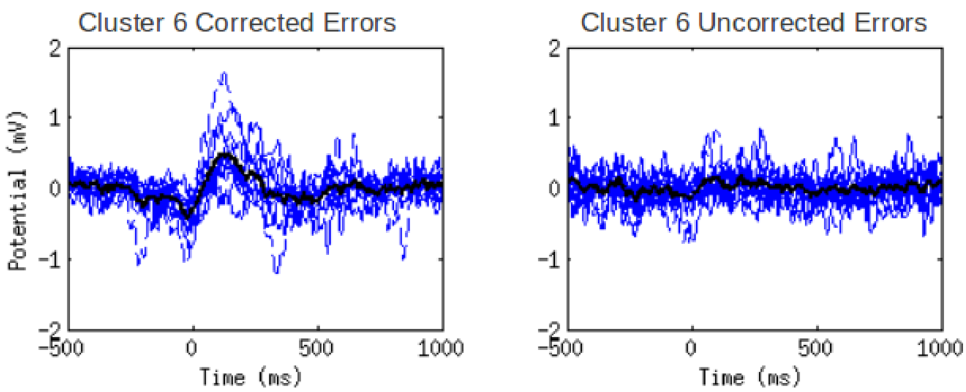


Figure 4.5: The average ERPs of the ErICs contributed by all participants during corrected error (left) and matched corrected (right) key-presses. Blue lines show average ERPs for each participant, and black line shows the grand average of the ERPs from all participants

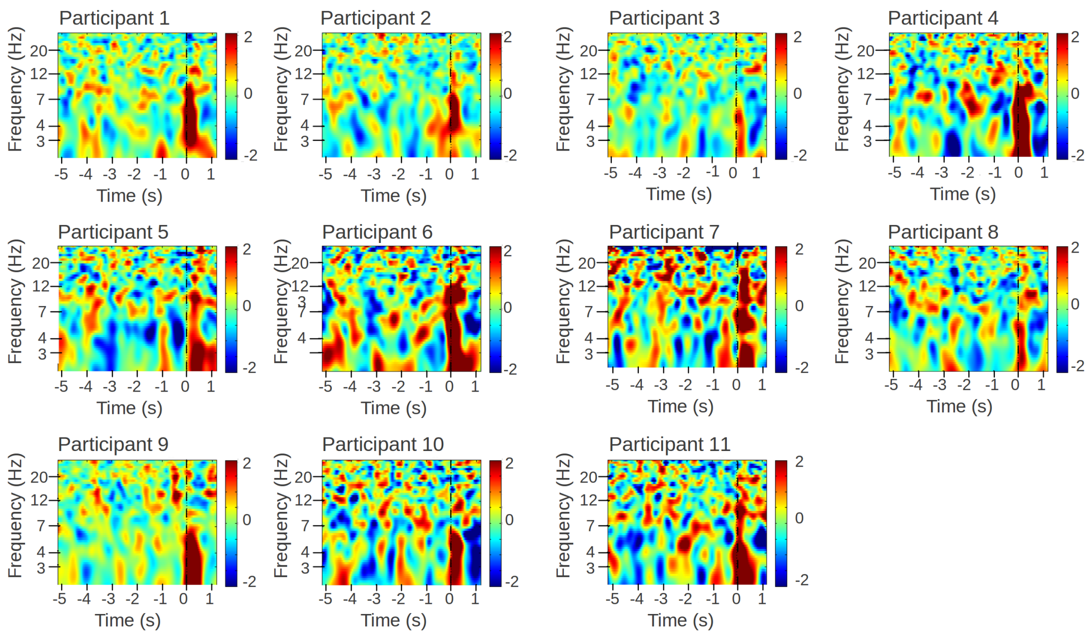


Figure 4.6: The average ERSP of the ErICs contributed by each participant during corrected error key-press. Vertical dashed line shows the time of key-press. Magnitude of ERSP represented by the colour scale.

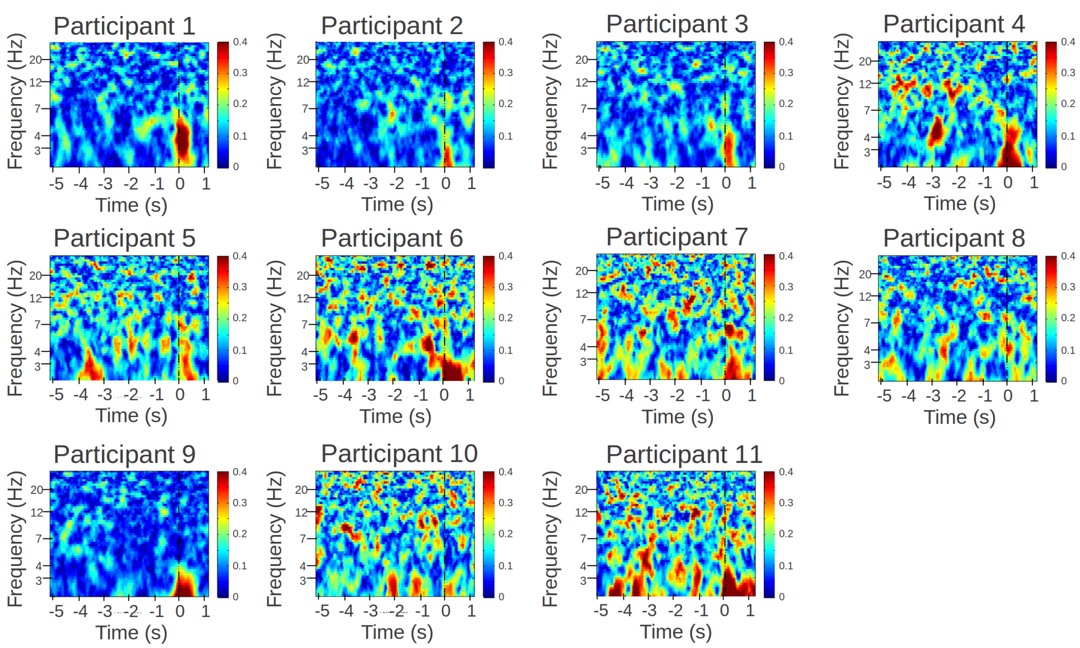


Figure 4.7: The average ITC of the ErICs contributed by each participant during corrected error key-press. Vertical dashed line shows the time of key-press. Extent of ITC represented by the colour scale.

peaked 19.53ms before the key-press, and Pe peaked 113.28ms after. Paired samples t-test showed that the ERP differences between corrected and matched correct key-presses were significant at times of ERN ($p < 0.01$) and Pe ($p < 0.01$) peaks.

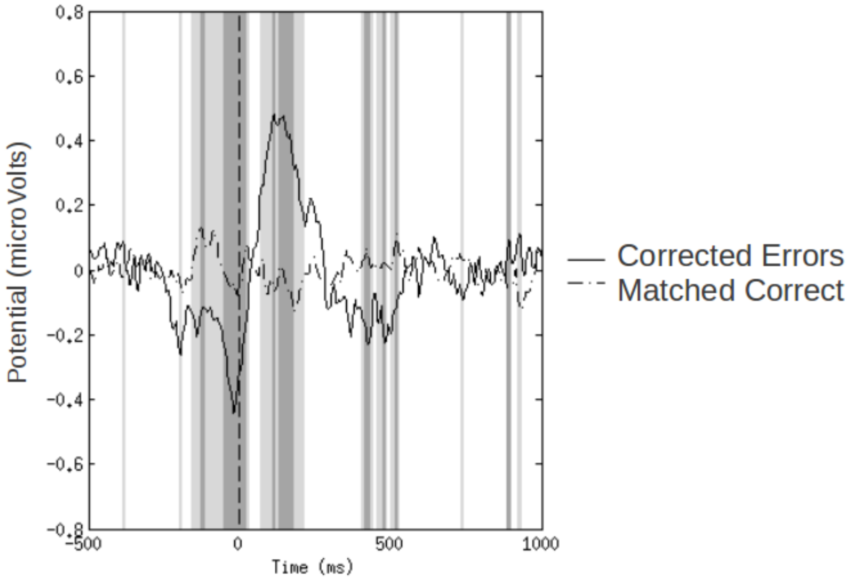


Figure 4.8: ERPs associated with corrected error presses and matched correct key-presses. Areas during which the two ERPs are significantly different are plotted in light ($p < 0.05$) an darker ($p < 0.01$) grey. Vertical dashed line shows the time of key-press.

We haven't used any corrections for multiple comparisons as we had strong *ad hoc* predictions about 1) the timing of the ERN and Pe (time of error key-press), 2) its polarity (negative), 3) its topography (fronto-medial areas), 4) sensitivity to errors (stronger association with errors than correct key-presses). Thus we were confident that the negative EEG component during the time of error key-press (but not the correct key-press), which was maximal over the fronto-medial parts of the scalp was the ERN and the positivity that shortly followed it was the Pe described in the literature. Because we had empirically based expectations as to when, where and how the ERN and Pe would appear *ad hoc*, we saw no necessity to apply corrections for multiple comparisons.

However we did find an unexpected difference about 340 to 530ms after the error key-press. Because this difference was not predicted by any previous work that we are aware of, we further tested its reliability. We conducted the same paired samples t-test on the amplitudes of the correct vs. detected error ERPs in the period 300 to 600ms after the response, this time applying corrections for multiple comparisons. As shown in figure 4.9, we found that the differences between corrected error ERP and matched correct ERP were not continuously reliable, slipping in and out of significance at 0.05 alpha level.

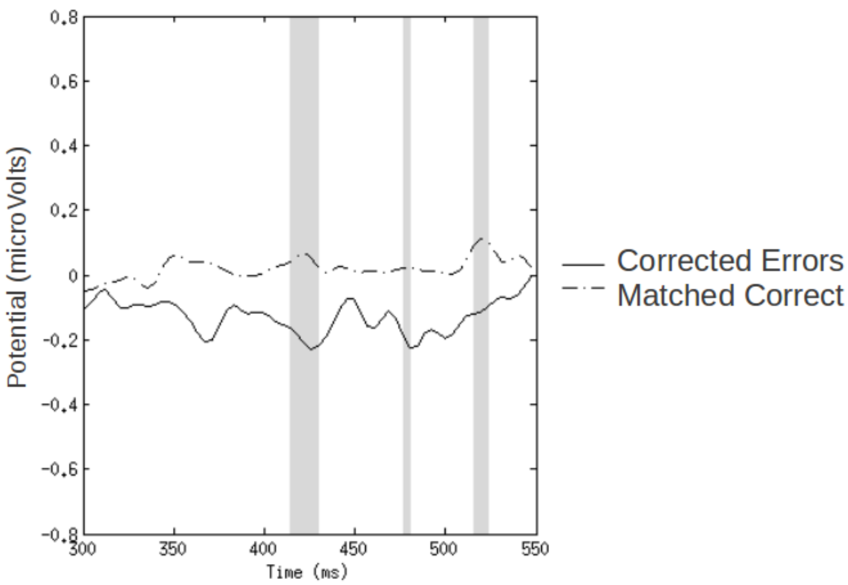


Figure 4.9: The difference between the corrected error ERPs and matched correct ERPs between 300 and 550ms after the key-press. Areas during which the two ERPs are significantly different are plotted in light grey ($p < 0.05$).

Uncorrected Error vs. Matched Correct Key-press ERPs

A comparison of the ERPs associated with uncorrected error key-presses to those associated with matched correct key-presses revealed no significant differences neither at the time of the key-press (no ERN), nor afterwards (no Pe). There were isolated time points where the difference between the uncorrected error key-presses and matched correct reached significance (figure 4.10), but because these were differences not predicted *ad hoc*, we followed them up using corrections for multiple comparisons. None of these differences remained significant after correction for multiple comparisons.

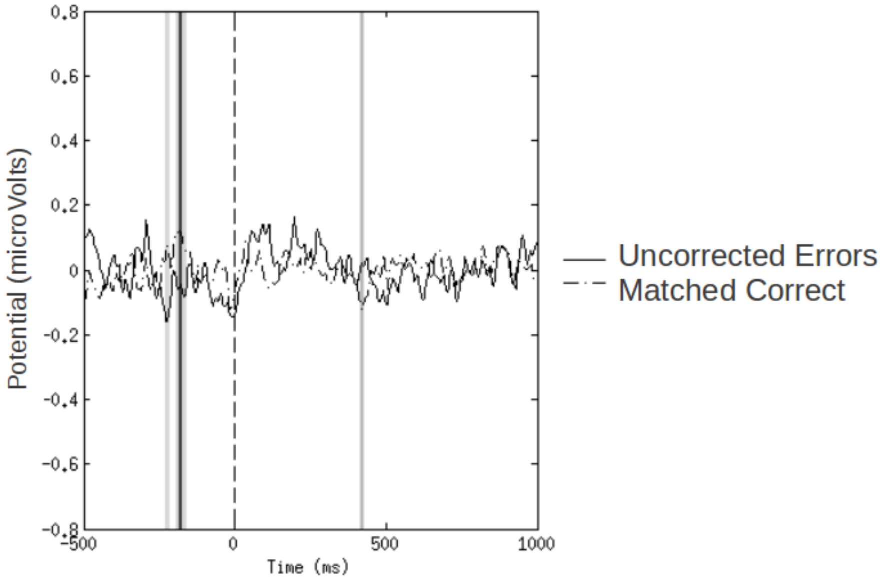


Figure 4.10: ERPs associated with uncorrected error presses and matched correct key-presses. Shaded areas highlight the time points when the uncorrected error and matched correct keys had significantly different activations. Vertical dashed line shows the time of key-press.

Corrected vs. Uncorrected Error Key-presses

A comparison of ERPs associated with corrected vs. uncorrected errors is shown in figure 4.11. We found that corrected errors were associated with stronger ERN and Pe components than uncorrected errors.

4.3.3 Correlations Between ERP and Behavioural measures

Before calculating the statistical correlations between ERP measures (i.e. ERN and Pe amplitude) and behavioural measures (i.e. error and post-error slowing), we plotted all error key-press instances by all participants, sorted according to post-error slowing (figures 4.12 & 4.13) and error slowing (figures 4.14 & 4.15). These plots show that there was no strong relationship between the amplitude or the latency of ERN or the Pe. In other words, error key-presses which were associated with high error or post-error slowing were not found to be associated with stronger or earlier ERN amplitude or latency, compared to those associated with low error

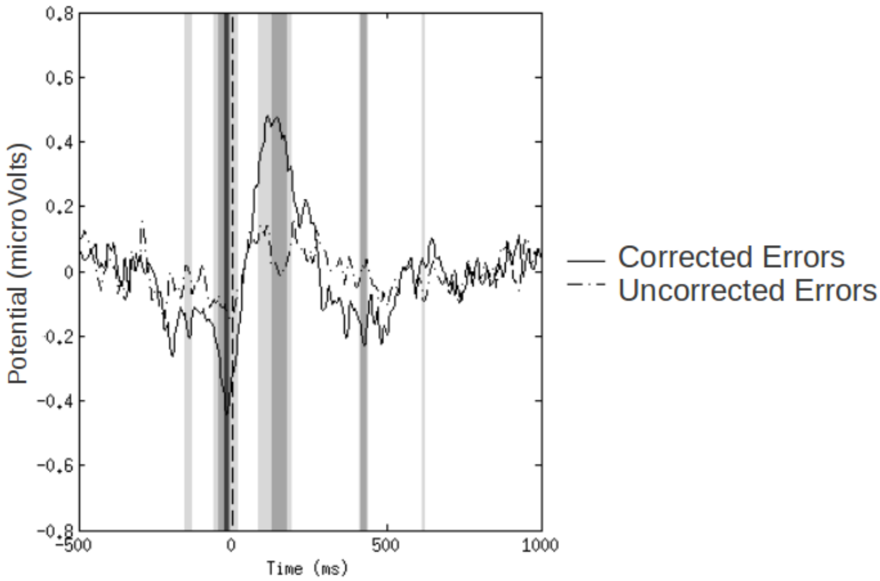


Figure 4.11: ERPs associated with the corrected and uncorrected errors. Areas during which the two ERPs are significantly different are plotted in light ($p < 0.05$), medium ($p < 0.01$) and darkest ($p < 0.001$) grey. Vertical dashed line shows the time of key-press.

or post-error slowing.

It is important to note here that the number of trials included in these plots are different than the actual number of trials in other analyses. This is the case particularly for figure 4.13. First, A total of more than 500 corrected errors were included in the analysis. However, vast majority of these were followed immediately by backspace, and thus have no post-error key-presses. This is the same reason we do not report post-error slowing in experiment 1 (chapter 3). There were simply not enough correct post-corrected-error key-presses. Second, there were no post-uncorrected-error key-presses for uncorrected errors which were typed at the end of the word. Third, the first letters of sentences do not have an IKI and thus error slowing values. Fourth, error letters inserted at the ends of words can't be matched for letter position and word length. For example, an extra $5th$ letter in a 4 letter word (e.g. 'word**d**') can't be matched with a correct letter because 'word' is a 4 letter word). We didn't calculate error slowing for errors which could not be matched in terms of letter position and word length. Fifth, we found that many uncorrected errors were not followed by correct key-presses, and thus many post-uncorrected error key-presses were also not possible to match to a correct key-press in word length and letter position. In other words, when an error key-presses is followed by another error key-press, we didn't calculate the post-error slowing for the second error key-press.

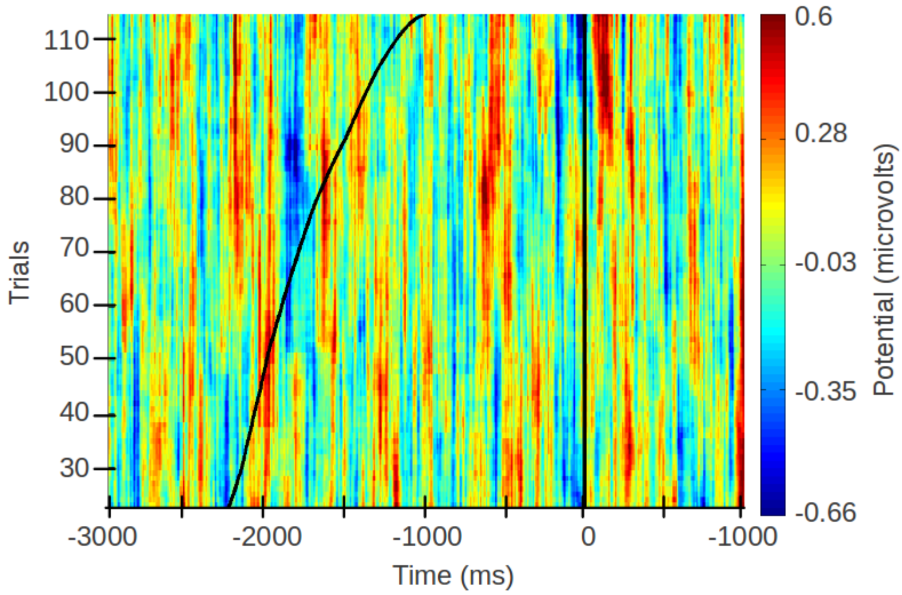


Figure 4.12: Event related changes in electrical potential (microVolts, as represented by the colour scale) during uncorrected errors sorted by post-error slowing. Trials associated with highest post-error slowing are plotted up and vice-versa. Solid vertical line shows the time of key-press. Curved line to the left shows the amount of post-error slowing, but is not referenced to the times on the x-axis (i.e. post-error slowing was not all negative).

Even though the plots show no clear relationship between behavioural and ERP effects of errors, we calculated correlations for each participants data. The investigation of across participant patterns of EEG-behaviour correlations (i.e. correlations between the EEG effects and behavioural effects of errors) was conducted in a manner identical to the investigation of behavioural effects such as error slowing.

First we calculated a Spearman's rho value for the EEG-behaviour correlations (e.g. the correlation between error slowing and ERN amplitude) for each participant. This gave us one rho value for each EEG-behaviour pair for each participant. Then we applied statistical tests to check if the null hypothesis (that the across participant average of these rho values are no different than 0) can be rejected. Figures 4.16 and 4.17 show the distributions of correlation coefficients contributed by the participants

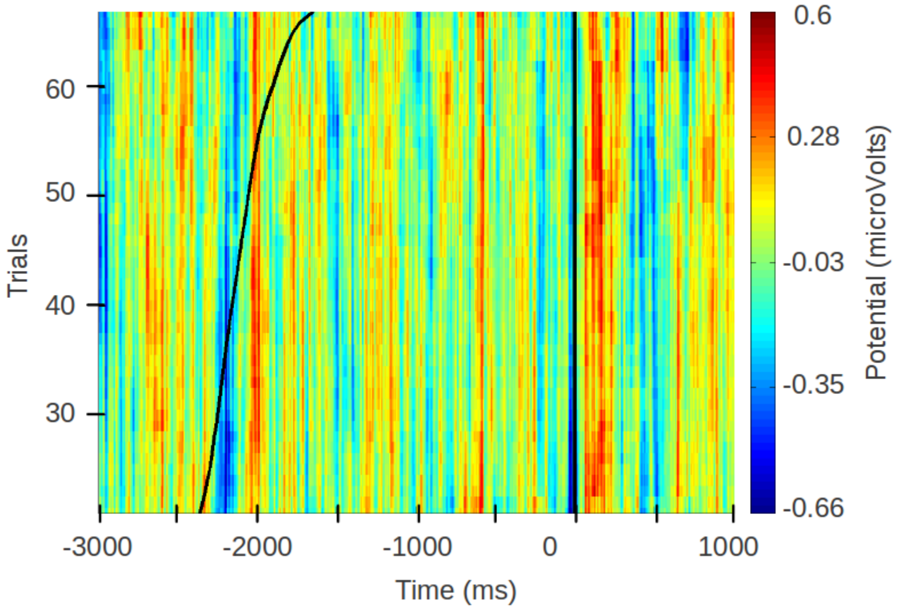


Figure 4.13: Event related changes in electrical potential (in microVolts, as represented by the colour scale) during corrected errors sorted by post-error slowing. Trials associated with highest post-error slowing are plotted up and vice-versa. Solid vertical line shows the time of key-press. Curved line to the left shows the amount of post-error slowing, but is not referenced to the times on the x-axis (i.e. post-error slowing was not all negative).

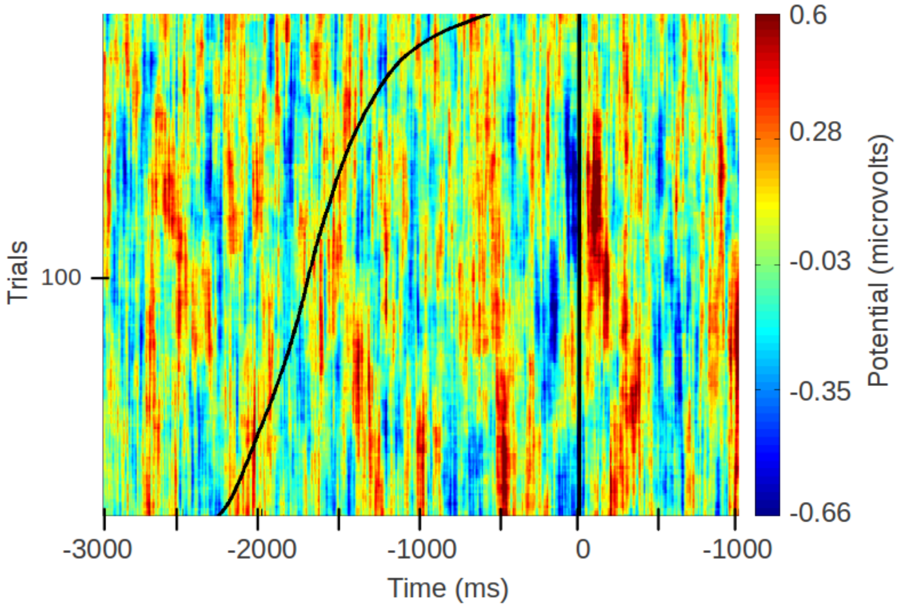


Figure 4.14: Event related changes in electrical potential (in microVolts, as represented by the colour scale) during uncorrected errors sorted by error slowing. Trials associated with highest error slowing are plotted up and vice-versa. Solid vertical line shows the time of key-press. Curved line to the left shows the amount of post-error slowing, but is not referenced to the times on the x-axis (i.e. post-error slowing was not all negative).

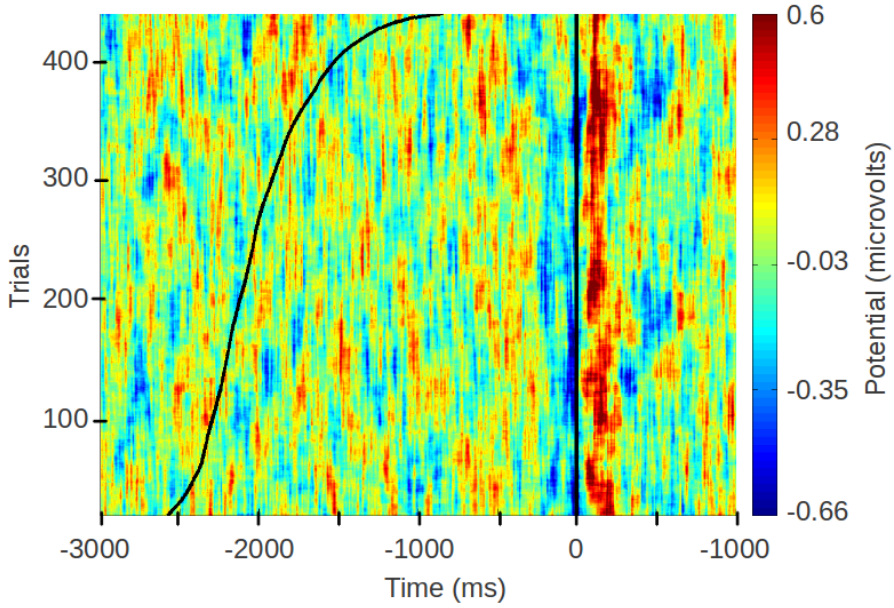


Figure 4.15: Event related changes in electrical potential (in microVolts, as represented by the colour scale) during corrected errors sorted by error slowing. Trials associated with highest error slowing are plotted up and vice-versa. Solid vertical line shows the time of key-press. Curved line to the left shows the amount of post-error slowing, but is not referenced to the times on the x-axis (i.e. post-error slowing was not all negative).

to the across participant average for the correlations. Neither error slowing nor post-error slowing tend to be correlated with neither Pe nor ERN amplitude or latency. Wilcoxon's signed rank tests showed that the mean of none of these histograms were different than zero (all p's > 0.083).

4.4 Discussion

Our aims in this experiment were to replicate and extend the error related ERP findings to skilled typing and then to study the relationship between ERN and Pe and error correction; and error slowing. In addition, we aimed to replicated the behavioural findings we report in the first experiment (i.e. chapter 3. Below we summarize our findings and discuss them in relation to findings in the literature.

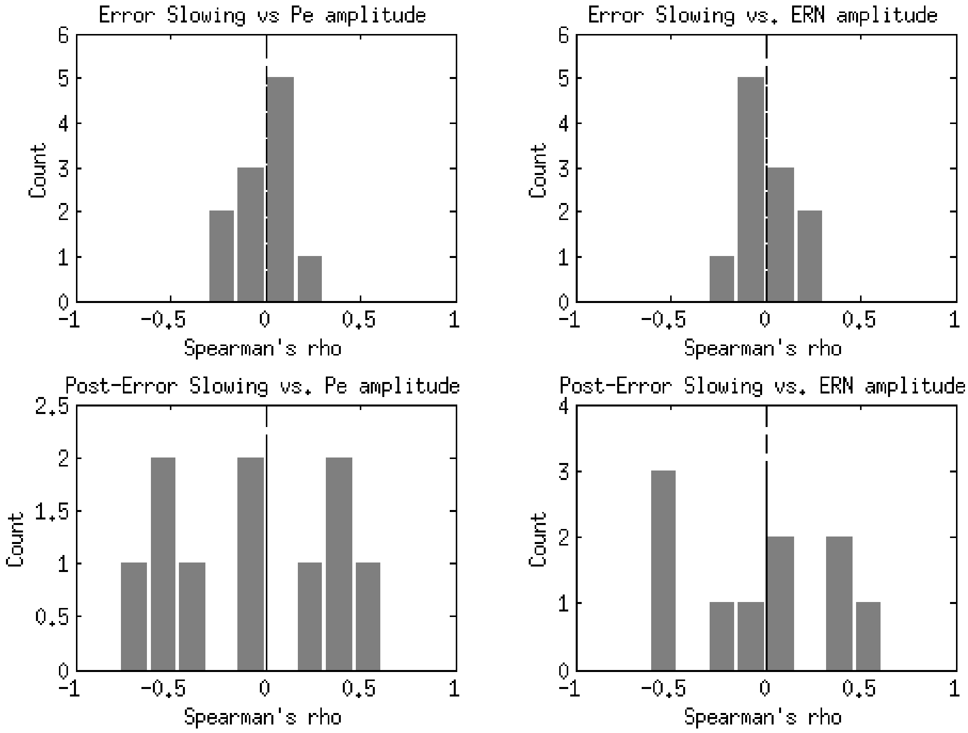


Figure 4.16: Histograms showing the distribution of correlation coefficients between EEG and behavioural measures. Upper left - Error Slowing vs. Pe amplitude; Upper right - Error slowing vs. ERN amplitude, Lower left - Post-Error Slowing vs. Pe amplitude, Lower right - Post-Error Slowing vs. ERN amplitude

4.4.1 Behavioural Findings

Our behavioural analyses showed that the typing speed is slowed down during and after the pressing of the wrong key. However, in the current experiment, we failed to find any pre-error speeding effects. This is in contrast to the pre-error speeding effect we observed in our first experiment.

Theoretical implications of error and post-error slowing are discussed in chapter 3, so will not be repeated there. However, failure to replicate the pre-error speeding in the current study warrants discussion at some length. It is possible that methodological differences between the first and the current experiments lead to contrasting findings of pre-error speeding. These are discussed below.

One obvious difference between the current experiment and the previous one was the length of the two experiments. Typists in the first experiment typed 20 sentences as opposed to 100 in the current one. One possibility is that the pre-error speeding is contaminated by fatigue caused by typing too many sentences. To check

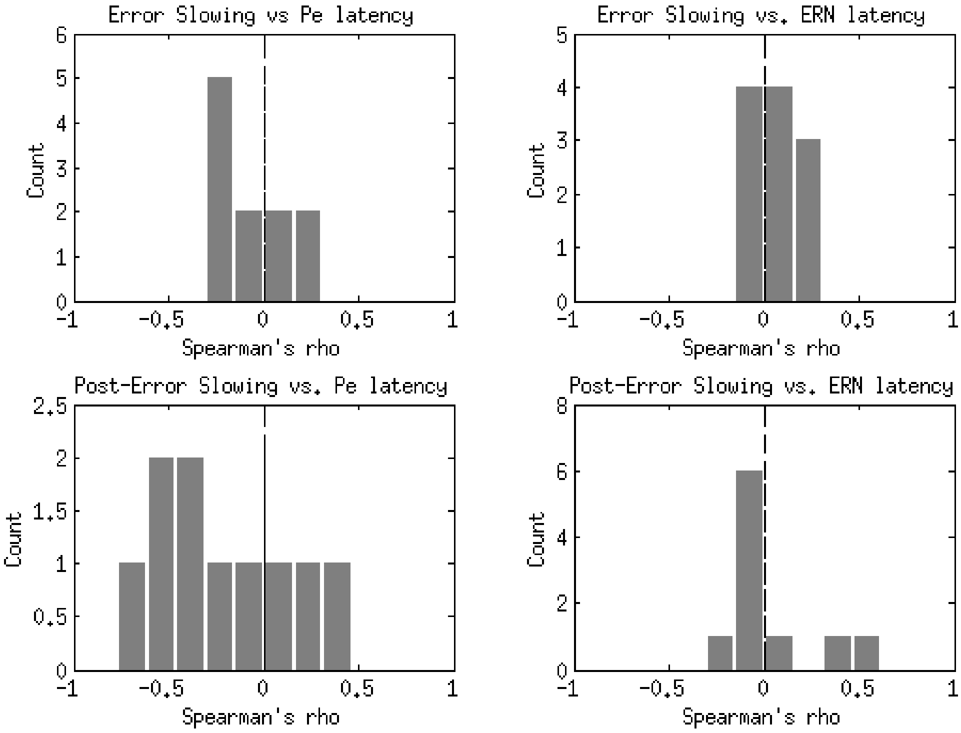


Figure 4.17: Histograms showing the distribution of correlation coefficients between EEG and behavioural measures. Upper left - Error Slowing vs. Pe Latency; Upper right - Error slowing vs. ERN Latency, Lower left - Post-Error Slowing vs. Pe Latency, Lower right - Post-Error Slowing vs. ERN Latency

if this was the case, we analysed the typing performance of the participants during the first 20 sentences. Even when we only looked at the first 20 sentences typed by participants, we found no significant pre-error effects. Looking at the next set of 20 sentences also revealed no pre-error speeding effects, whereas error and post-error slowing were evident in all sections of the data. Because we found no pre-error speed differences between the beginning and the end of typing performance, we concluded that the lack of pre-error speeding in the current experiment was not due to typing too many sentences.

Another difference was in the instructions used in the second experiment. The first experiment revealed that there was a large variability on the error correction rate of the participants. For example, there was one participant who corrected all of her errors (i.e. no detected errors) and another participants who didn't miss any of her errors (i.e. no undetected errors). This created problems in terms of the number of observations for each key-press class (as discussed in section 3.2.3). In order to avoid this issue in the second experiment, the importance of correcting all errors made was heavily emphasized by the experimenter, during verbal instructions. Unfortunately, we can not compare the effect of this difference in instructions on pre-error speeding, because all of the participants started the experiment with the same instructions.

Other differences between the two experiments include different participants, sentences and the involvement of EEG recording. However, none of these differences affected other measures of error performance (i.e. the error and post-error slowing). For these reasons, rather than explaining away the lack of pre-error speeding by differences between our experiments, we acknowledge the possibility that pre-error effects are not as reliable and robust as the error or post-error slowing effects. Any small pre-error speeding effect might have been inflated due to the small number of observations we had in the previous experiment. Further research can reveal under what circumstances pre-error speeding effects are observed, but this is beyond the scope of this chapter.

4.4.2 EEG Findings

We have found that corrected error key-presses were associated with strong ERN and Pe components compared not only to matched correct key-presses but also to uncorrected error key-presses. Further, uncorrected errors were found to be no different to matched correct key-presses in terms of ERN and Pe. These results suggest both ERN and Pe might be related to error awareness. However, we found no relationship between these EEG measures and behavioural measures that have been previously associated with error detection (i.e. error or post-error slowing).

In the sections that follow we argue that our results point to EEG measures as more reliable indices of error awareness and error detection than the behavioural ones.

Measures of Error Detection

In our design, we had three potential markers of performance monitoring, all of which have previously been shown to be sensitive to errors of performance. These are the EEG measures (ERN and Pe), performance slowing measures (error and post-error slowing), and the pressing of backspace (an error signalling response).

Behavioural measures traditionally used as indices of performance monitoring include error signalling responses (Rabbitt, 1967, 1968; Shaffer & Hardwick, 1969; Ullsperger & von Cramon, 2006; Ursu et al., 2009), as well as post-error and error slowing (G. Logan & Crump, 2010; Rabbitt, 1978; Shaffer, 1975). The relationship between error signalling response and performance is straightforward in that when a participant makes an error signalling response the experimenter can be confident that the error was detected (i.e. the performance monitoring processes were engaged). The relationship between post-error slowing and performance monitoring however is more ambiguous. As shown in the two studies reported in this thesis and also by G. Logan and Crump (2010), even the errors which the participant is not aware of are followed by post-error slowing. Further, as also mentioned in the introduction chapter, the amount of post-error slowing is usually greater than necessary (Rabbitt & Rodgers, 1977) and is not always associated with an increase in post-error accuracy (Hajcak et al., 2003, 2005; Notebaert et al., 2009).

Thus there seems to be a lack of consistent relationship between the amount of post-error slowing and adaptive changes in post-error performance. Two emerging patterns from our data suggest that this is true in typing as well. First, we found that the EEG measures of performance monitoring and error signalling responses were in line. The errors which were corrected were associated with large ERN and Pe components, and those which were not corrected were similar to the correct keypresses in terms of ERN and Pe. Second, we found that error slowing and post-error slowing were present both in corrected and uncorrected errors (i.e. slowing measures were not as sensitive to error correction as the EEG measures). Third, we found no correlations between the EEG measures of ERN and Pe latency and amplitude and slowing measures: The errors which were associated with large ERNs were no more likely to be followed by increased error or post-error slowing.

All of the three points raised above are based on the assumption that pressing of backspace is a more valid marker of performance monitoring than error or post-error slowing. We believe that the validity of a measure of performance monitoring (or of any other cognitive skill) is defined heavily by the outcome or the final impact of that measure on performance. In the case of copy-typing, the task of the typist is to get a text written as accurately as possible. The effect of post-error slowing on typing performance is not as adaptive (e.g. post-error slowing after uncorrected errors) as that of pressing of the backspace, which eliminates the effect of error from the performance: Put another way, errors followed by a backspace during typing are much more likely to lead to an improvement in performance than those followed by post-error slowing. For this reason we believe that error correction via backspace is a much better measure to use when studying performance monitoring than error or post-error slowing.

4.4.3 Alternative Explanations

Lack of ERN during Uncorrected Errors

Many studies of ERN have established that detected errors are associated with large ERN and Pe amplitudes, whereas undetected errors are associated with smaller ERN amplitudes and no Pe. If one assumes that Pe reflects a neural process that is related to error awareness (Endrass et al., 2007; Nieuwenhuis et al., 2001; OConnell et al., 2007) and ERN a different process such as response conflict (Botvinick et al., 2001; Carter et al., 1998), then one may expect them to be represented by different ICs.

It is possible that in the current study i) using only the corrected error epochs and ii) specifying ITC as a weighted parameter in the clustering process lead to the grouping together of the ICs which showed both Pe and ERN as opposed to those showing only a weak ERN in the error related cluster. In other words, it is possible that the error related cluster included ICs related to error awareness but excluded those related to response conflict which leads only to ERN. This may have lead to the exclusion of ICs representing undetected error effects (such as response conflict) from the error related cluster. This would explain why we didn't observe any ERN during uncorrected errors.

Another observation that further concerned us was the observation of an unexpected common property of the ErICs in the error related cluster. A look at the figure 4.3 e) shows that the ICs in the error related cluster were associated with significant ITC at a lower than expected frequency range (i.e. 2-4Hz). This corresponds to the frequency of the error related changes following corrected errors. The effect of corrected errors on EEG (i.e. ERN + Pe) lasted for about 366ms (from -156ms to 210ms, see figure 4.10) which corresponds to a frequency of about 3Hz. This means that the ICs selected by the clustering algorithm underwent error induced phase locking in oscillations with wavelengths ranging from 250 to 500ms (i.e. frequencies ranging from 2 to 4Hz). In contrast, uncorrected errors are associated with no Pe, and thus have a shorter wavelength (wavelength was 175ms for ERN in our study) and a higher frequency (about 6Hz in our study), well within the theta range.

Because the clustering algorithm grouped ICs based on their ITC activations (among other properties) at the time of a *corrected* error key-press, it might have left out the ICs which showed ITC changes at higher frequencies (i.e. narrower wavelengths such as that of the ERN).

Even though we were concerned about the above points, there are a number of observations from and checks we applied to our analyses, supported by previous findings in the literature, which suggested that the lack of ERN during uncorrected errors was not caused by biased clustering of ICs. These are summarized below.

First, the available evidence from the EEG literature suggests that there is only one IC that represents both the ERN and Pe (Gentsch et al., 2009; Hoffmann & Falkenstein, 2010; Roger et al., 2010; Wessel et al., 2012). We are not aware of any ERN studies which decomposed EEG data using ICA and found separate ICs representing ERN and Pe. Thus we believe that the ICs selected using corrected error epochs would be the only ICs related to error related changes.

Second, ITC values that were taken into account by the clustering algorithm are within the frequency range of 4-8Hz. In other words, the clustering algorithm was blind to ITC changes taking place below 4Hz and above 8Hz. This suggests that the error induced ITC increase in 2-4Hz frequencies was an emergent quality of the ErICs in the error cluster, rather than a property that was used/detected by the cluster algorithm to group them together.

Third, even though errors lead to no significant ITC increases in frequencies higher than 4Hz, we observed error induced ERSP changes in all participants in oscillations of frequencies between 4 and 8Hz (see figure 4.6). Importantly, this suggested that corrected errors were indeed associated with higher frequency oscillatory changes, and our clustering algorithm was sensitive to these changes. Put another way, if there were ITC changes associated with higher frequency oscillations (as expected during an uncorrected error leading to an ERN but no Pe), our clustering algorithm would be able to detect and use them in grouping ICs together.

Fourth, to ultimately decide whether the observed lack of ERN during uncorrected errors was due to the clustering of ICs, or due to use of ICs (i.e. loss of data through ICA), we calculated the grand average ERPs from 4 fronto-medial electrodes (namely C20, C21, C22 and C23, see figure 4.18), time locked to error key-presses. We found that the same pattern of results were reflected by the fronto-medial electrodes (see figures 4.20 and 4.19). We concluded from these figures that the lack of a reliable ERN during uncorrected errors was not caused by a self selection of ICs which are only sensitive to corrected errors. The contrast between corrected error related changes and uncorrected error related changes was as evident in the electrode data as in the ErIC data.

Lack of Correlation between Error Slowing and EEG measures

The lack of observed correlation between the effect of errors on EEG and behaviour could be due to a number of reasons. First, it might be because there is truly no trial-by-trial correlation between EEG and behavioural error effects. A second possibility is that our measures were not sensitive to this correlation.

Remember that the data we used in calculating single trial correlations between

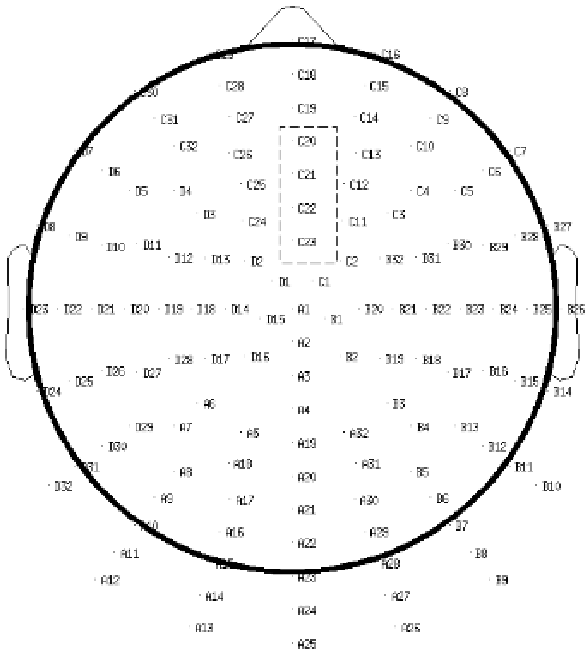


Figure 4.18: Locations of the electrodes C20, C21, C22 and C23, marked by the dashed rectangle

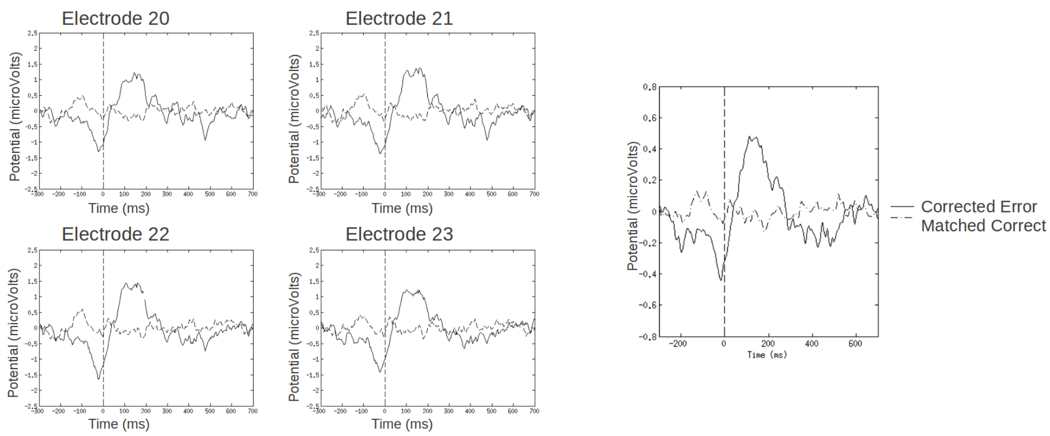


Figure 4.19: Event related potentials as recorded by fronto-medial electrodes (4 smaller figures on the left), and the activations in the Error related Cluster (bigger plot on the right) at the time a corrected error. Dashed vertical line shows the time of key-press.

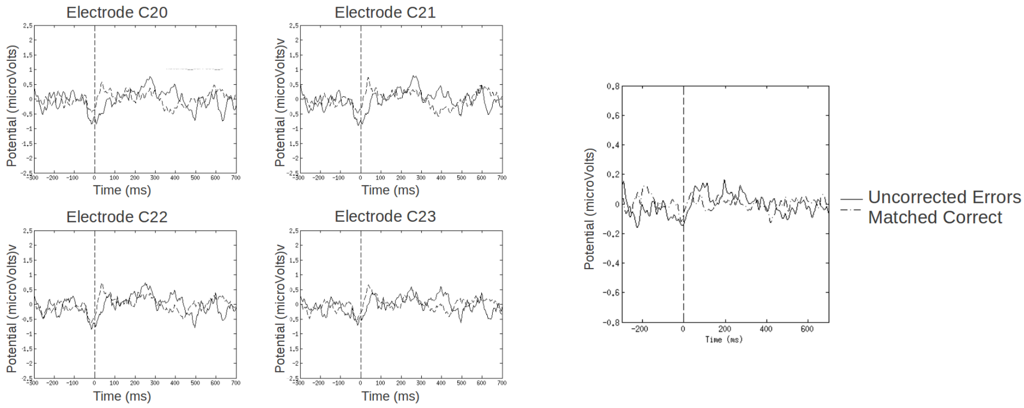


Figure 4.20: Event related potentials as recorded by fronto-medial electrodes (4 smaller figures on the left), and the activations in the Error related Cluster (bigger plot on the right) at the time an uncorrected error. Dashed vertical line shows the time of key-press.

behavioural slowing measures and EEG measures were the amount of error slowing and the peak EEG amplitude in a given time window (-156 to 19ms). Calculation of the amount of error slowing for a given error response is straight forward. This is how much slower an error key-press is compared to matched correct key-presses. On the other hand, the definition of error effects on EEG is more ambiguous, particularly because of the lack of a well-established time window in which the error effects are expected to appear.

As a simple check to see if the time window we used was effective in revealing the effect of errors on single trial EEG record, we collected the most negative EEG amplitudes within the window for each corrected and uncorrected error (i.e. negative peak amplitude) and called it single-trial ERN peak (or stERN). Then, we compared the average of stERNs for corrected errors to that recorded during uncorrected errors. Because we previously found a strong difference between the grand average ERPs of corrected and uncorrected errors, we expected the average of stERNs recorded within this window to be significantly different between corrected and uncorrected errors too. However, we found that the average of the stERNs in this window was no more negative during corrected errors than uncorrected errors. This suggests strongly that either the *time window* we used, or the measure of single trial *amplitude* was not sensitive to effect of errors on EEG.

To check if this was due to the measure of single trial amplitude we used, we tried the same analysis using 3 other different single trial amplitude measures, keeping the time window constant (from -156 to 19ms after the key-press). These were

i) the average EEG amplitude during -156 to 19ms after the key-press,

ii) the difference between the most negative peak during -156ms to 19ms after the response and the most positive peak during 258ms and 156ms before the response, and

iii) the difference between the average amplitude during 156ms before and 19ms after the error and 258ms before and 156ms before the error key-press.

None of these measures were effective in revealing a difference between corrected and uncorrected error amplitudes within the time window -156ms to 19ms after the error key-press. As a result we concluded that the time window we specified and/or the single trial peak amplitude measures we used were not sensitive enough to the effects of errors on EEG.

As a consequence, we refrain from making any strong claims about the single trial relationship between the EEG and behavioural effects of errors: The lack of correlation between the ERN/Pe and error slowing was likely to be caused by the insensitivity of the single trial measures we used. More research is needed to establish the critical time window within which the effects of errors on EEG become clear, but this was not one of the primary aims of the current study.

4.4.4 Performance Monitoring and EEG measures

In short, our results can be summarized in these points: 1) EEG components ERN and Pe reflect a single/unitary process, which plays a key-role in performance monitoring and error correction: The errors which were eventually corrected were preceded by a large amplitude ERN and Pe, and those which were not corrected were preceded by no such EEG changes. This difference in ERN and Pe amplitude between corrected and uncorrected errors is a strong and statistically reliable one. 2) The amplitude or the latency of these measures do not seem to be related to the amount of error or post-error slowing associated with errors. In fact, these EEG measures can be dissociated from error and post-error slowing in that ERN and Pe are not observed after uncorrected errors, but both error and post-error slowing are.

Outstanding Questions We show in this experiment that ERN and Pe are both electro-physiological markers related to performance monitoring and error detection/correction. However, certain important questions related to the essence of these cognitive processes remain unanswered. One such question is related to the temporal dynamics of the process that gives rise to the ERN and Pe. It is not clear to us from the current analyses or from the current literature whether the ERN and Pe are the manifestations of

i) An online oscillatory mechanism which continually monitors the performance - such that when it is 'engaged', errors are detected and when it is not, errors are not detected. The only way this would be detected using ERP analysis would be if this oscillation under-went phase locking in response to the error. Increased power in the oscillations (as an indication of engagement), would not necessarily show on the average ERP as they would be averaged out unless accompanied by phase locking.

ii) A reactive system which becomes engaged in response to an error. If this engagement is strong enough (above a threshold), errors are detected, and if it isn't, errors are not detected. By definition this process suggest that changes in the EEG are triggered by the error response. This would mean that the EEG effects would be phase locked to the error key-press. This would be reflected in the ERPs. Further, the EEG components that appear in the grand average ERP would reflect the physical characteristics of the underlying oscillations (i.e. its wavelength and frequency). A look at figures 4.8 and 4.10, suggest to us that i) when the reaction of this system is beyond a certain threshold, errors are corrected (figure 4.8), and when it is silent errors go uncorrected (figure 4.10) and ii) the neural oscillations underlying this system seem to have a wavelength of about 366ms (i.e. a frequency of about 3Hz).

However, because we haven't done any statistical analyses of EEG data in the frequency domain, the points raised above can not go far beyond speculation. In the following chapter, we present results from the time-frequency analysis of the same EEG dataset, to further investigate relationship between performance monitoring and the EEG measures.

4.4.5 Conclusions

We show for the first time that both ERN and Pe are strongly related to error detection and performance monitoring and improvement in a real world action. We further show that error slowing and post-error slowing are not necessarily adaptive in improving the performance as they are associated with corrected as well as uncorrected errors. This contrast between the relationship of EEG measures and slowing measures to adaptive performance changes (i.e. pressing the backspace) is in line with the lack of correlations between the error slowing and the EEG measures. Thus we conclude this chapter by claiming that the ERN and Pe are important indices of performance monitoring and improving whereas error and post-error slowing are post-error behavioural reactions, which may only partially reflect performance monitoring mechanisms.

Chapter 5

Experiment 2: Frequency Domain Analyses of Errors

5.1 Introduction

As mentioned earlier in section 1.2.4, EEG oscillations are found to be involved in a number of cognitive tasks. The aim of this chapter is to focus on the fronto-medial theta oscillations in the context of performance monitoring. We will do this by extracting a number of parameters (i.e., event related spectral perturbations, inter-trial coherence) from the ICs in the error related cluster and studying how they change in response to errors (error vs. matched correct keys) and error correction (corrected vs. uncorrected errors).

In the sections that follow, we summarize the current knowledge on the relationship between the theta parameters and errors and performance monitoring. Then, we state our predictions and hypotheses based on this knowledge and go on to describe the analyses we conducted to test them.

5.1.1 Error Responses and Theta Power and Synchrony

Error responses lead to increases in power and synchronization in theta frequency oscillations. One of the earlier reports of consistent effects of errors on EEG oscillations was published by Luu and Tucker (2001). These authors asked participants to do a flankers task and replicated the observation of ERN following error responses. These authors showed that ERN could be a partial manifestation of an ongoing oscillation of theta frequency range by applying a band-pass filter to isolate 4Hz-12Hz oscillations. A later study by Luu et al. (2004) showed quite convincingly that theta oscillations and ERN are strongly aligned in time, using single trial as well as grand average EEG traces. It is clear from the single trial results of (Luu et al., 2004, figures 2 to 5) that the phase of the theta oscillation was reset to more or less the same angle after an error, irrespective of its pre-error phase. This realignment of the phase of a given oscillation in response to an external event is referred to as phase-locking or entrainment (Buzsaki, 2006). Even though it has been shown that ERN can be generated without phase-locking in theta oscillations using simulated EEG data (Yeung, Bogacz, Holroyd, Nieuwenhuis, & Cohen, 2004), a number of studies using actual EEG data and different statistical and analytical methods since then replicated the finding that increase in power of and partial phase-locking in theta oscillations underlie the appearance of ERN (Cavanagh, Zambrano-Vazquez, & Allen, 2011; Trujillo & Allen, 2007), and predict post-error changes to performance (Cavanagh et al., 2009; Cohen & van Gaal, 2012; Herrojo-Ruiz et al., 2011). Trujillo and Allen (2007)'s report is particularly convincing as these authors outlined 3 possible mechanisms by which the ERN can be formed and showed that both phase resetting (i.e. ITC) and power increases (i.e. ERSP) in theta frequency oscillations are important factors for the generation of ERN.

5.1.2 Current Analyses

The literature summarized above is dominated by studies using discrete trial tasks (except for Herrojo-Ruiz et al. (2011)), which did not consider error awareness (except for Cohen, van Gaal, Ridderinkhof, and Lamme (2009)). We are not aware of any reports which studied performance monitoring using an ecological tasks where error awareness and post-error adjustments are studied in relation to EEG oscillatory parameters such as theta band activity.

In the current analyses, we aimed to contribute the current literature by using a continuous action, putting a strong emphasis on the relationship between to theta oscillations and i) error awareness and ii) error related performance changes (e.g. error slowing and error correction). Further we were interested in any potential differences between detected vs. undetected errors, *before the error response*. Based on previous work we predicted that errors would lead to greater theta power and inter-trial phase coherence than matched correct key-presses. Further, if these parameters are associated with adaptive changes to behaviour, they should also be more likely to be followed by backspace.

The motivation for studying pre-error period was related to the essence of performance monitoring system. If performance monitoring is sub-served by an online neural process, then the pre-error state of the oscillations representing this process should predict the probability of detecting an error. If on the other hand, performance monitoring is sub-served by neural processes which are sensitive to errors, pre-error state of the oscillations representing it will not be predictive of error detection. Rather, the magnitude of the response of this system to the error event will be predictive of error detection. Thus, by studying the pre-error state of a range of oscillations, we can make inferences about the essence of the neural mechanisms sub-serving performance monitoring system.

5.2 Methods

The data presented here come from the same experiment as that reported in chapter 4. Thus the participants, design, and procedure was identical to the first experiment. Only the steps followed in the extraction of frequency domain from the EEG data are described in the below paragraphs.

Time-frequency transformations were accomplished using the built-in functions of EEGLAB (i.e. newtimef and newcrossf). The mathematical basis for these functions are described in detail in Delorme and Makeig (2004, pp. 14-17). The parameters used for accomplishing the time-frequency transformations were the default ones in the EEGLAB, as detailed below.

5.2.1 Epoch Length

We used epochs from 6 seconds before to 2 seconds after the key-press. As we mentioned earlier, the motivation behind selecting such a wide epoch window was to investigate the state of these oscillatory processes before the error. This way, we could check if changes in certain oscillations before the error moderate error detection.

5.2.2 Event Related Spectral Perturbations (ERSP) and Inter-trial Coherence (ITC)

ERSP

The newtimef function of EEGLAB was used to extract power at each frequency. Power (in decibels, dB) is calculated by computing the power spectrum over a sliding latency window using wavelets at each frequency and time point. Baseline normalization is achieved by subtracting log power spectrum during the baseline period from the spectral power at each time point and frequency (Delorme & Makeig, 2004). Statistical tests are then applied to check if there are reliable ERSP changes associated with different key-presses. The ERSP and ITC analyses were conducted only for ErICs in the error related cluster.

Baseline Period The baseline period used to calculate the average log power spectrum at each frequency was 6000ms to 200ms before the key-press.

Frequency Range The ERSP was calculated for 57 logarithmically spaced frequencies from 2Hz to 30Hz. Even though we expected error related changes mainly in the theta (4-8Hz) frequency range, we kept the frequency range wider for exploratory purposes.

Wavelet Cycles and Window Length The number of wavelet cycles in the time-frequency transformations used was 3 (with a sliding window length of 1668ms) at the lowest frequency (i.e. 2Hz), increasing to 22.5 in the highest (i.e. 30Hz) in linear steps.

ITC

The ITC analyses were conducted in the same manner and using the parameters as the ERSP ones: We used the phase values returned by the newtimef function in ITC analysis. ITC is a measure of temporal relationship between the phase/angle of a given oscillation and the onset of an event (e.g. a typing error): An ITC of 1 suggests that the phase of the oscillation at the time of the event for all different trials was identical (perfect time-locking between the onset of the event and the phase resetting of the oscillation), and an ITC of 0 suggests the phase of the oscillation was random (phase of the oscillation not affected by the event at all).

Statistical tests were then applied to check whether ITC values associated with different kind of key-press (i.e. errors and matched correct key-presses) were reliably different. All of these calculations were conducted within EEGLAB environment.

5.2.3 Single-trial ERSP and Error Slowing Correlations

As emphasized many times in previous chapters, we believe it is important to study the links between the effects of errors on EEG and behavioural measures. To investigate the relationship between the error related ERSP changes and performance slowing, we extracted 4 measures from each error key-press epoch: Error slowing, post-error slowing, peak theta ERSP amplitude (at 4Hz between 97.66ms before to 464.84ms after the error key-press), and ERSP peak latency (i.e. the time at which the ERSP amplitude reached its maximum between 97.66ms before to 464.84ms after the error key-press). The time window of -97.66 to 464.84ms was determined based on the plots presented in figure 5.1. Briefly, -97.66ms was the first time point in 4Hz oscillations when the ERSP associated with corrected errors was statistically and continuously different than that with matched correct key-presses, and 464.84ms was the last.

After extracting these measures for each epoch, we calculated single trial correlations between EEG measures (peak theta ERSP and latency) and behavioural slowing measures (error and post-error slowing), to see if the epochs associated with strong and timely theta ERSP bursts were more likely to be associated with strong error slowing and vice-versa.

5.3 Results

5.3.1 Corrected Errors vs. Matched Correct Key-presses

Error Related Spectral Power

Corrected error key-presses were found to be associated with significantly stronger ERSP values than matched correct key-presses. As shown in figure 5.1, there was a strong increase in ERSP with an onset of 97.56ms, 148.44ms and 128.91ms *before* the error key-press at 4Hz, 6Hz, and 8Hz, respectively. These are the times at which the difference between the ERSP associated with corrected errors and matched correct key-presses was reliable at an alpha level of 0.05, not corrected for multiple comparisons (see figure 5.1c, for results after correction for multiple comparisons). The offset of this difference was at and 464.84ms 332.03ms and 566.41ms for 4Hz, 6Hz and 8Hz oscillations, respectively.

We also observed a strong decrease in power at oscillations of 4.28 to 5.69Hz between 808.59 and 1074.22ms after the corrected error key-presses, which was statistically reliable after correcting for multiple comparisons.

Inter-trial Coherence

Similarly, ITC values associated with corrected errors were found to be significantly higher than those associated with matched correct key-presses. The ITC values for corrected errors were already significantly stronger than those for matched correct key-presses 62.50ms before the key-press. This difference persisted until 316.41ms after the key-press, for 4Hz oscillations. No significant ITC effects were observed for higher theta oscillations such as the 6Hz or 8Hz. This was in contrast to significant ITC at lower frequency oscillations (see figure 5.1).

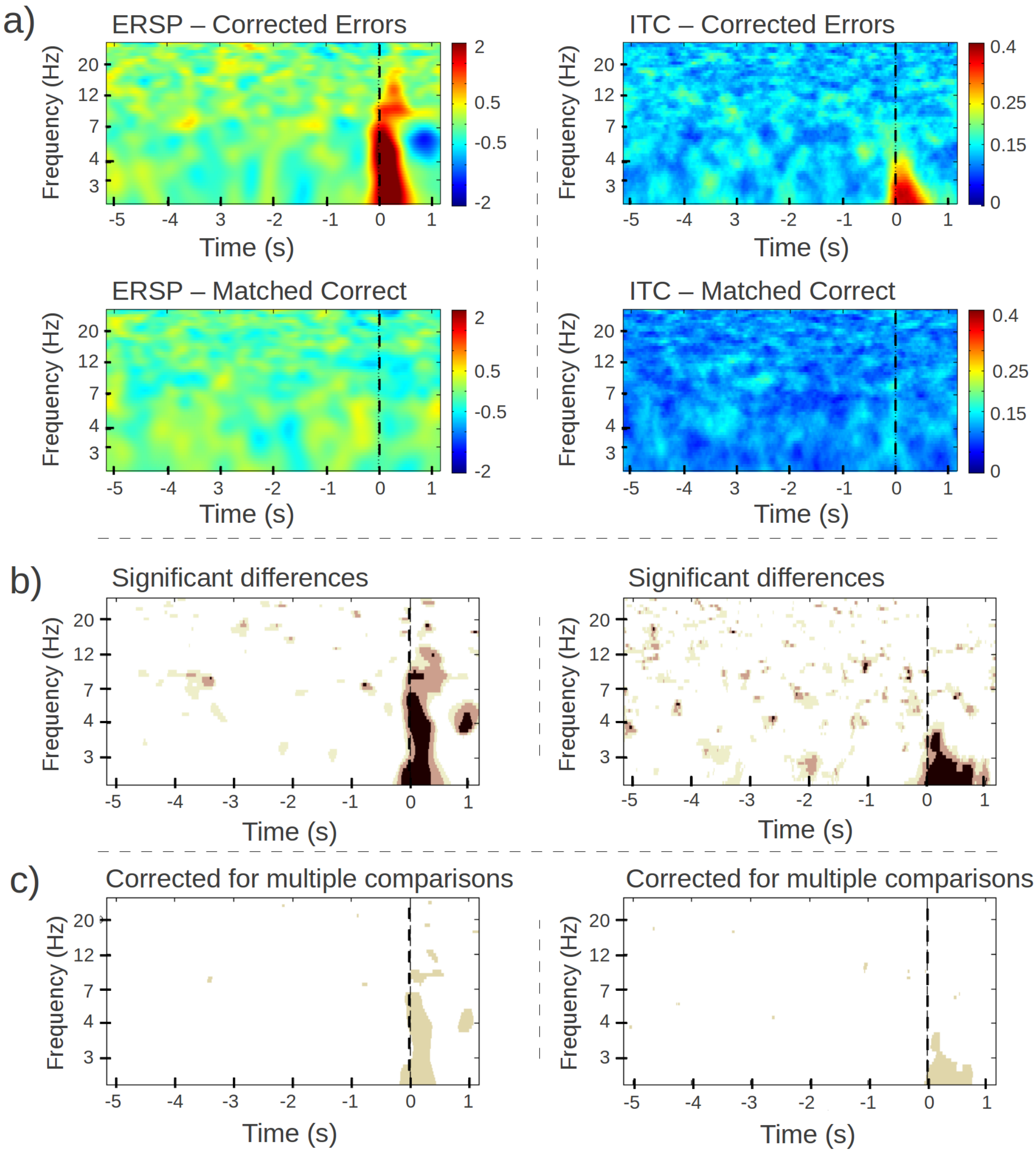


Figure 5.1: a) Average ERSP (left column) and ITC (right column) values for corrected error key-presses (first row) and matched correct key-presses (second row). The magnitude of ERSP and ITC are represented by the colour scale. b) Statistical significance of the difference between the two key-presses (white - $p > 0.05$, light brown - $p < 0.05$, medium brown - $p < 0.01$, dark brown - $p < 0.001$), before correction for multiple comparisons. c) Statistical significance after correcting for multiple comparisons at $p < 0.05$. Dashed vertical line shows the time of key-press.

5.3.2 Uncorrected Errors vs. Matched Correct Key-presses

Error Related Spectral Power

As shown in figure 5.2, uncorrected errors were also associated with increased ERSP values compared to matched correct key-presses. The average ERSP values during uncorrected error key-presses were larger than those during matched correct key-presses from 121.09ms before to 480.47ms after the key-press at 4Hz oscillations, and from 93.75ms to 230.47ms after the key-press at 6Hz oscillations. However, none of these differences were statistically significant after correcting of multiple comparisons.

Inter-trial Coherence

Uncorrected error key-presses were not associated with statistically stronger ITC values than matched correct key-presses (figure 5.2).

5.3.3 Corrected vs. Uncorrected Error Key-presses

Error Related Spectral Power

As shown in figure 5.3, corrected errors were associated with stronger ERSP values than uncorrected errors. The onset of this difference was 70.31ms and 117.79ms before the key-press at 4Hz and 6Hz oscillations, respectively. These differences lasted until 273.44ms and 316.41ms after the key-press, for 4Hz and 6Hz oscillations. Oscillations at 8Hz were only different between 70.31ms and 152.34ms after the key-press.

Inter-trial Coherence

We found strong differences between corrected and uncorrected errors ITC values. As shown in figure 5.3, corrected errors were associated with higher ITC values than uncorrected errors from -54.69ms to 242.19ms after the key-press at 4Hz oscillations and even earlier for lower frequency oscillations (e.g. from -74.22ms to 359.38ms at 3Hz). However, this pattern of ITC differences was not observed in higher theta

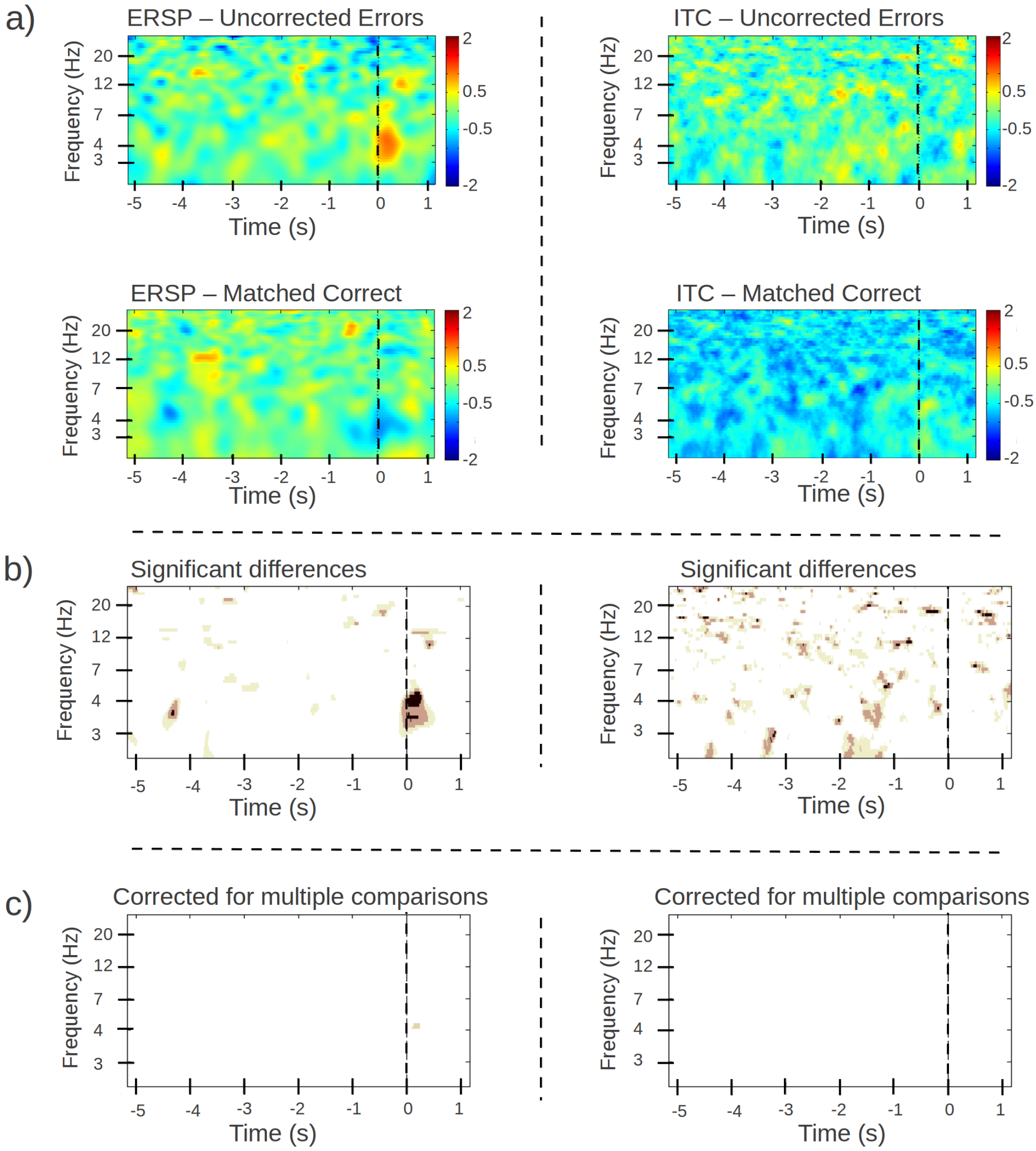


Figure 5.2: a) Average ERSP (left column) and ITC (right column) values for uncorrected error key-presses (first row) and matched correct key-presses (second row). The magnitude of ERSP and ITC are represented by the colour scale. b) Statistical significance of the difference between the two key-presses (white - $p > 0.05$, light brown - $p < 0.05$, medium brown - $p < 0.01$, dark brown - $p < 0.001$), before correction for multiple comparisons. c) Statistical significance after correcting for multiple comparisons at $p < 0.05$. Dashed vertical line shows the time of key-press.

frequency range. Further, none of these differences were found to be reliable after correction for multiple comparisons.

5.3.4 Single Trial Correlations Between Theta Power and Behavioural Measures

Figure 5.4 shows the ERSP amplitudes associated with corrected and uncorrected error key-presses, sorted according to the amount of error slowing and post-error slowing for that key-press, across all participants and all error key-presses. These plots show no strong relationship between the amount of error or post error slowing and ERSP amplitudes during corrected error key-presses, but a possible link between ERSP and the amount of error slowing in uncorrected errors. The plot on the bottom left shows that the uncorrected errors which were associated with the highest amount of error slowing were also associated with high ERSP amplitudes.

To check if this pattern was a statistically reliable one, we calculated the trial by trial correlations between peak ERSP amplitude and error slowing for each participant. The reliability of across participant average correlations was determined and statistically checked in the same manner as described in section 4.3.3. Briefly, we extracted one correlation coefficient (e.g. spearman's rho associated with theta peak latency vs. error slowing) value from each participant, representing the strength and direction of relationship between the ERSP and error slowing. Then, across participant average of these correlation coefficients were checked to see if they were significantly different than 0 using one sample t-tests.

We found a weak correlation between the error related ERSP magnitude and error slowing (across participant average for Spearman's rho= 0.145) for uncorrected errors. A one sample t-test showed that this was not statistically reliable at an alpha of 0.05 ($t(10) = 1.95$, $p = 0.080$). However, we found a similarly weak but more reliable relationship between the *peak latency* of ERSP and error slowing. The average correlation between peak ERSP onset and error slowing was -0.165. A one sample t-test showed that this mean was significantly different than 0 ($t(10) = -3.72$, $p = 0.004$). As such, the uncorrected error key-presses which were associated with

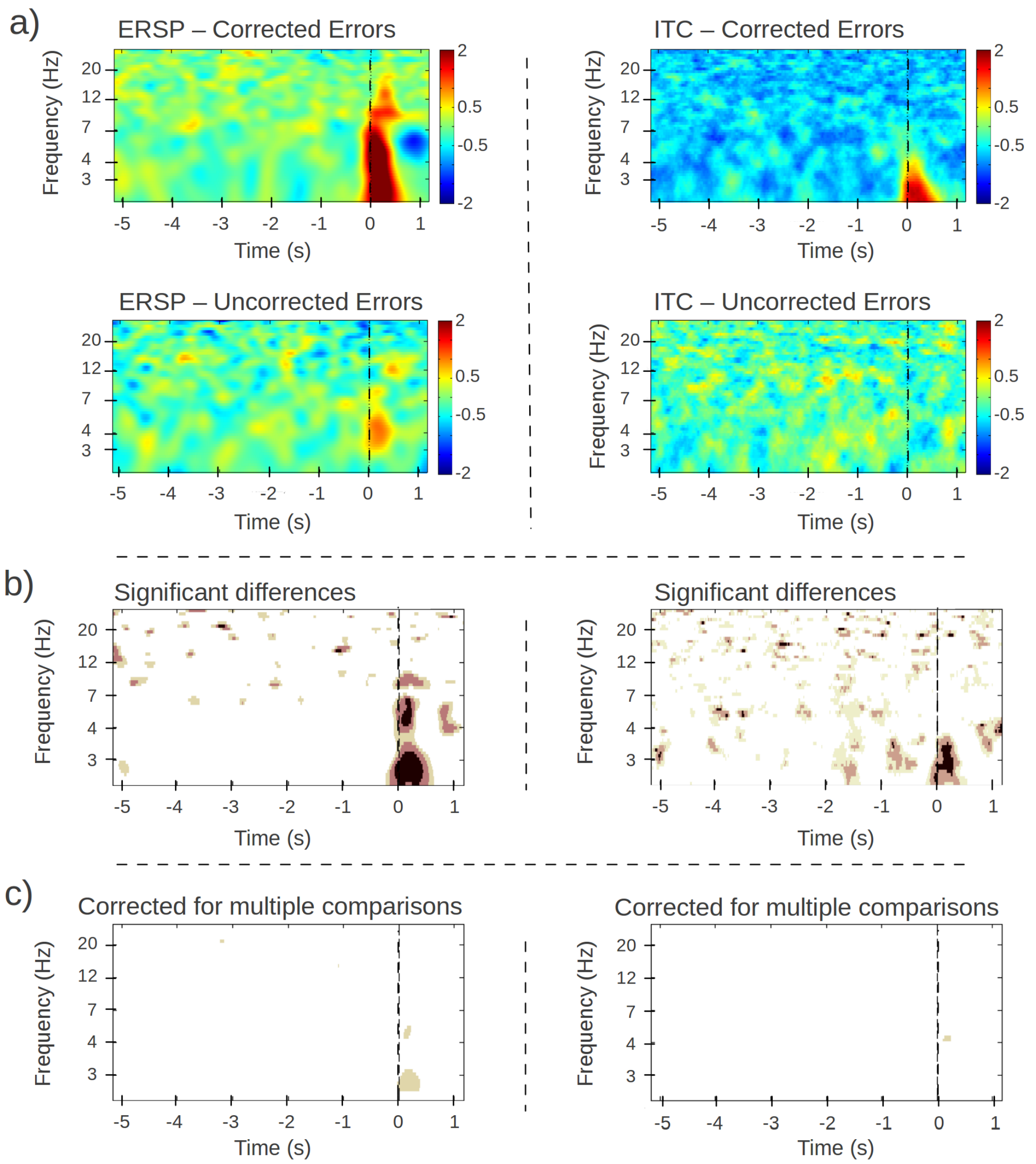


Figure 5.3: a) Average ERSP (left column) and ITC (right column) values for corrected (first row) and uncorrected error (second row) key-presses. The magnitude of ERSP and ITC are represented by the colour scale. b) Statistical significance of the difference between the two key-presses (white - $p > 0.05$, light brown - $p < 0.05$, medium brown - $p < 0.01$, dark brown - $p < 0.001$), before correction for multiple comparisons. c) Statistical significance after correcting for multiple comparisons at $p < 0.05$. Dashed vertical line shows the time of key-press.

earlier onset of ERSP peak were also associated with greater error slowing.

We found no such reliable correlations between ERSP amplitude or latency and post-error slowing. Similarly, none of the ERSP and slowing effects of errors were reliable for corrected error key-presses. Figures 5.5 and 5.6 show the distribution of correlation coefficients which gave rise to the means reported above.

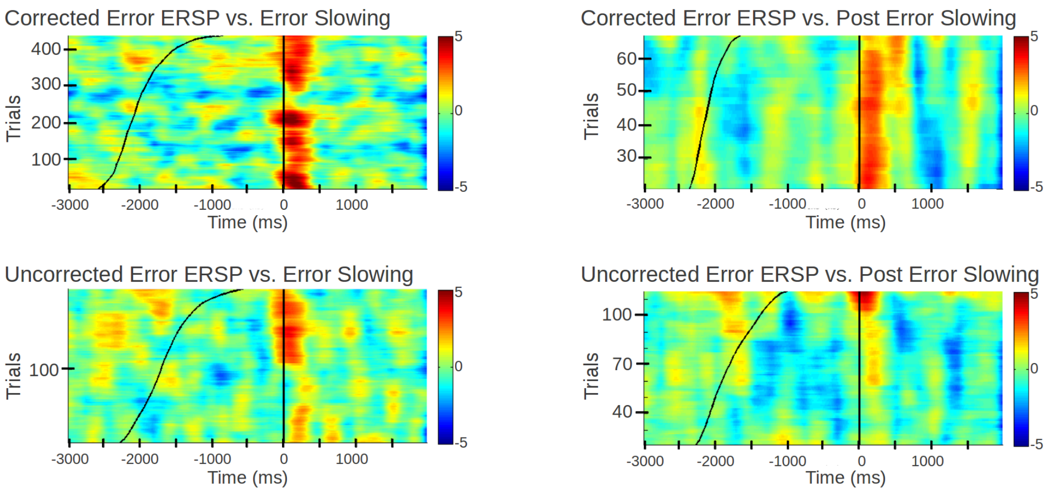


Figure 5.4: Changes in theta (4Hz) ERSP (dB, as represented by the colour scale) in response to corrected errors (top row) and uncorrected errors (bottom row). Vertical line shows the time of error key-press, and the curved line to the left shows the amount of error (left column) and post-error (right column) slowing associated with every trial. The trials are sorted according to the amount of slowing such that those associated with the highest amount of slowing are plotted at the top of each plot, and vice versa. The line on the left showing the amount of slowing for each trial is *not* referenced to the times on the x-axis (i.e. slowing values are not all negative).

As a reference, we also provide below histograms showing the relationship between correct key-press IKIs and the theta power at the time of these key-presses. Figure 5.7 shows this relationship for correct keys matched to corrected errors and figure 5.8 for those matched to uncorrected errors.

Finally, to check if the relationship between spectral perturbations at the theta frequency range and error correction we observed at the group level was reliable at a participant level we plotted figure 5.9. We found that the relationship between theta power and error correction probability was strong at the participant level. Figure 5.9 shows the pattern of relationship between peak ERSP and error correction probability for each participant. We divided all error key-presses according to the strength of peak theta power at the time of error key-press into bins (x-axis shows

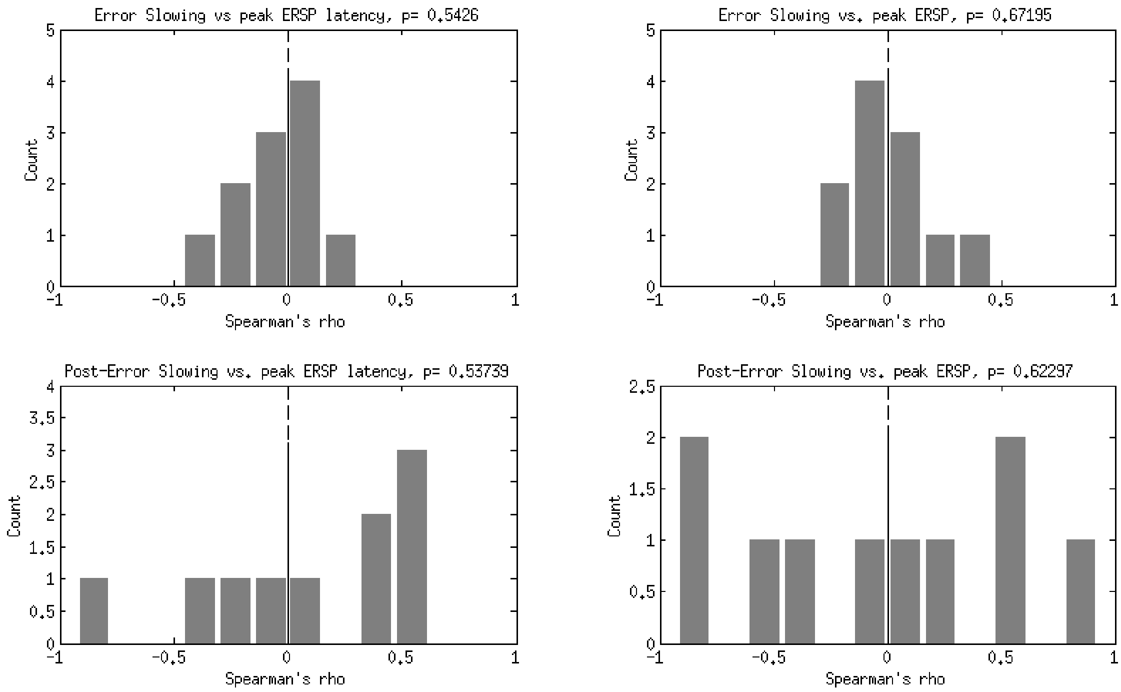


Figure 5.5: Distribution of correlation coefficients for ERSP and slowing measures corrected errors. Top left, error slowing vs. peak ERSP latency; top right, error slowing vs. peak ERSP amplitude; bottom left, post-error slowing and peak ERSP latency; bottom right, post-error slowing vs. peak ERSP amplitude. The p-values reported on top of the figure show the probability that the mean of the distribution is different than 0.

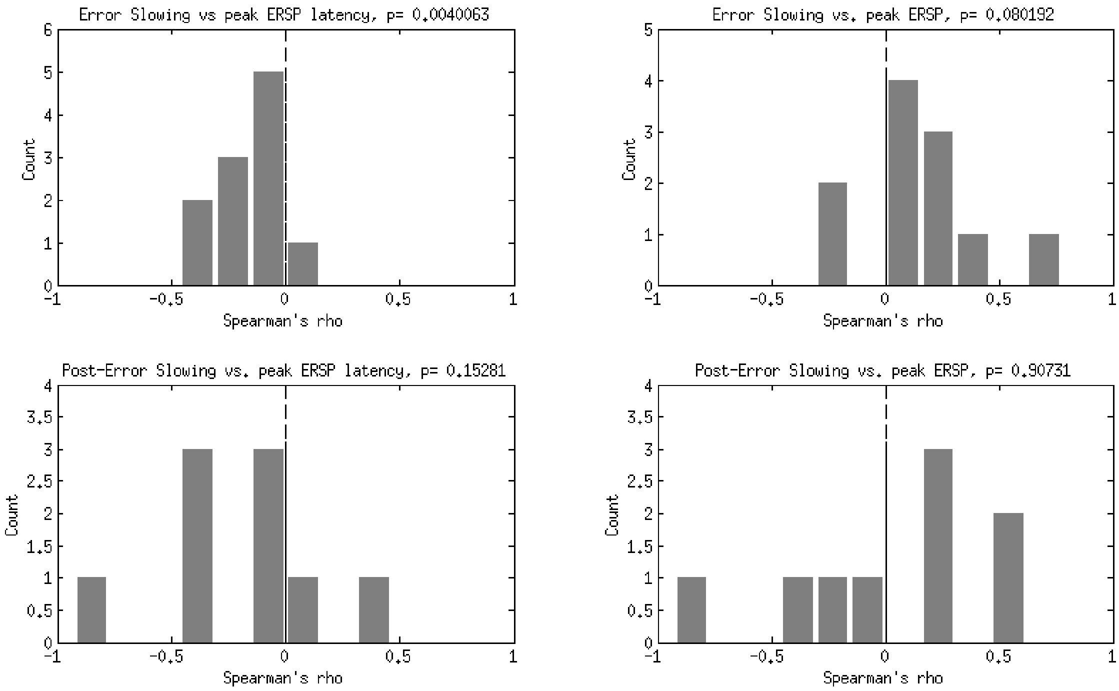


Figure 5.6: Distribution of correlation coefficients for ERSP and slowing measures during uncorrected errors. Top left, error slowing vs. peak ERSP latency; top right, error slowing vs. peak ERSP amplitude; bottom left, post-error slowing and peak ERSP latency; bottom right, post-error slowing vs. peak ERSP amplitude. The p-values reported on top of the figure show the probability that the mean of the distribution is different than 0.

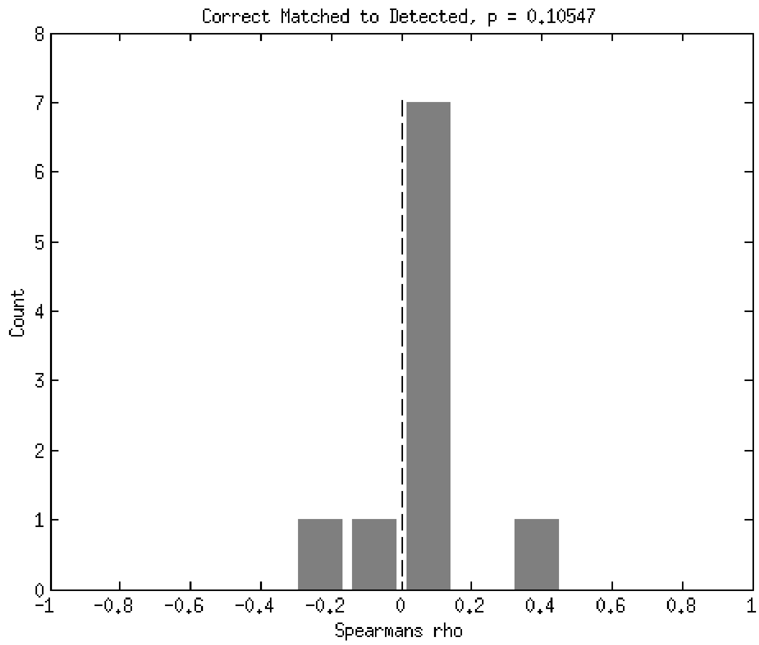


Figure 5.7: Histogram showing the correlation coefficient between the amplitude of the peak theta power during the correct key-presses matched to corrected errors and their IKIs. The p values reported in the title of the figure shown the probability that the mean of the coefficients is different than zero, based on Wilcoxon's signed ranks test.

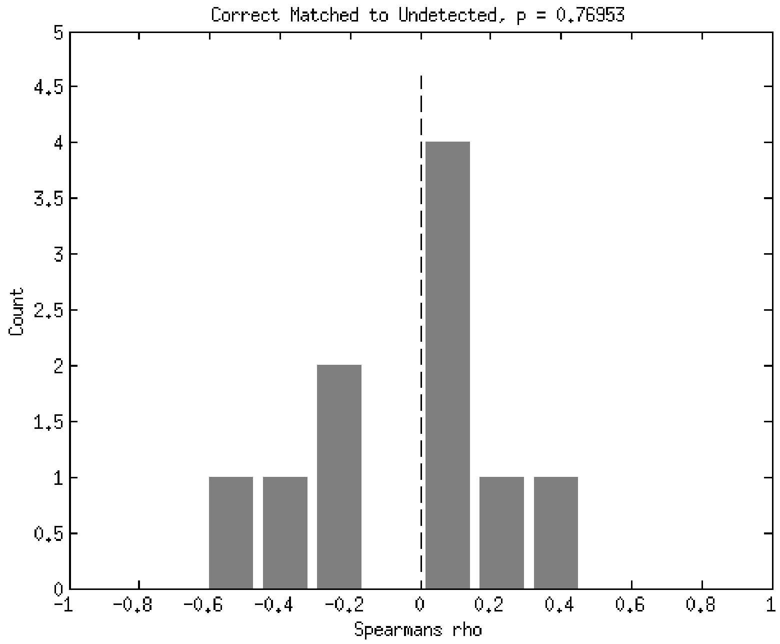


Figure 5.8: Histogram showing the correlation coefficient between the amplitude of the peak theta power during the correct key-presses matched to uncorrected errors and their IKIs. The p values reported in the title of the figure shown the probability that the mean of the coefficients is different than zero, based on Wilcoxon's signed ranks test.

the standardized peak theta power). Then we looked at the proportion of errors corrected to all errors, for each bin (y-axis shows the proportion of corrected errors). Finally, we calculated the correlation between theta power and the proportion of corrected errors, for each participant.

The mean correlation coefficient for all participants was 0.69. A Wilcoxon signed rank test showed that this positive correlation was a significantly reliable one ($p < 0.001$).

5.4 Discussion

The aim of the current set of analyses was to examine the effect of errors on EEG theta oscillatory activations as described in the *Current Analyses* section. We summarize and discuss the effects of errors on different parameters of EEG oscillations in the sections that follow.

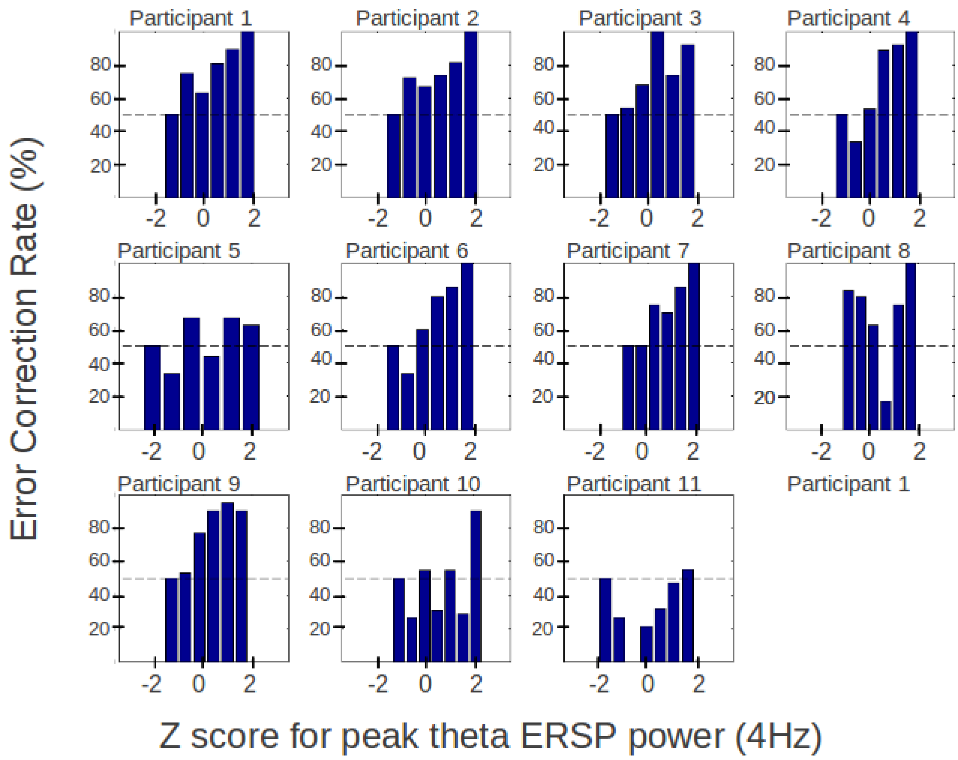


Figure 5.9: Ratio of corrected errors to all errors committed by each participant (y axis) as a function of the peak ERSP at 4Hz. The ERSPs are converted to z scores for each participant to improve comparability across participants. Horizontal dashed line shows 50% chance level

5.4.1 Pre-Error Oscillatory Activity

We found no differences between the corrected and uncorrected errors during the 6 seconds before the error, in the range of frequencies included in the current analysis (i.e. between 2 and 30Hz). This null finding has important implications for a claim we made in chapter 3 under the heading *Outstanding Questions*. Briefly, we speculated that pre-error mental states can be predictive of error correction probability, such that errors committed while the participants are 'in the zone' of typing will be more likely to be detected than those committed while they are not. However, we find that there are no reliable pre-error differences in the EEG parameters used in the current analysis. This issue will be discussed in greater length in the general discussion, where more across-chapter discussions are presented.

5.4.2 Theta Oscillations and Error Correction

Theta ERSP and Error Correction

We found that the greatest amount of theta power increase was associated with corrected errors, followed by uncorrected errors, and least with matched correct key-presses. Despite this strong relationship between the theta power and error correction, we found no such relationships between theta related parameters (peak theta power or latency) and error and post-error slowing. The only exception to this pattern was the significant correlation between the latency of theta power peak and error slowing during uncorrected errors (figure 5.6).

This association between peak theta power latency and error slowing in uncorrected errors was unexpected. However, a closer look at the lower left image in figure 5.4 shows that this pattern was driven by a simple effect. Uncorrected errors with above average error slowing were almost always preceded by a theta burst, whereas those with below average error slowing were either followed by a late (post-response) and weak theta burst or none at all. Even though we did not make such a prediction, this observation is sensible in hind-sight: If the theta power bursts are associated with error correction or awareness, they can affect the error key-press only if they precede it. In fact, figure 5.6 suggests that if a theta burst is present, it almost always precedes the key-press. Uncorrected error responses not preceded by theta bursts are associated with smaller error slowing than those preceded by theta bursts. To re-iterate, if the theta burst doesn't precede the error, it is either considerably smaller or not present at all.

We believe the observations summarized above suggest that theta oscillations are indeed intimately involved in error correction processes rather than 'error processing' per-se. As outlined in the introduction chapter, inferring error detection from error correction is usually problematic because some objectively uncorrected errors may well be subjectively detected or unsure errors. Figure 5.5 shows that the corrected errors were preceded by very strong theta power bursts but the amount of error slowing or post-error slowing was not related to the onset or the peak magnitude of theta power. However, as shown in figure 5.6, this was not the case for uncorrected errors: Even though some of the uncorrected errors were associated with theta power bursts, they were all associated with above average error slowing. Uncorrected errors which were not associated with theta bursts on the other hand, were associated with below average error slowing.

Looking at the pattern of theta activity and behavioural parameters, our interpretation favours a close link between theta power and error correction. It is not error actions themselves that are associated with theta activations, but the subjective experience of errors that lead to theta bursts.

A first look at the average theta power across different conditions suggests that there is a continuous relationship between theta power and error correction such that the higher the theta power, the higher the error correction rate (i.e. theta power at corrected errors > theta power at uncorrected errors > theta power at correct key-presses). We checked this assumption by plotting the error correction rate as a function of peak theta ERSP for each participant (figure 5.9).

Figure 5.9 shows that for most of the participants, the higher the ERSP, the greater the proportion of corrected errors to all errors. Based on these findings, we conclude that ERSP at theta range oscillations are strong predictors of performance monitoring and error awareness processes. They are also closely associated with adaptive changes in ongoing performance.

Theta ITC and Error Correction

We found that ITC in theta oscillations were also related to error correction responses. We found that corrected errors were associated with strong ITC at theta oscillations whereas uncorrected errors were not.

An indirect observation also shows that strong ITC in theta oscillations was present in corrected errors. Figure 5.1 shows a strong and significant decrease in power at 5Hz oscillations after the initial burst. We are not aware of such negative ERSP results in the performance monitoring literature, however, Trujillo and Allen (2007, figure 1b) shows mathematically and with modelling work that one manifestation of strong phase resetting in the absence of equally strong increase in power is a subsequent decrease in power.

Observation of such a strong ITC effect was not an effect we anticipated for the reasons outlined below. First, theta oscillations are not sensitive to motor processes but the accuracy of motor responses. If it is the error awareness that leads to phase-locking of theta oscillations, and if errors are slowed, then the onset of error response (the time locking event in the current analysis) will be shifted away in time from the onset of error awareness. Consequently, the phase/angle of the theta oscillations will vary with the amount of error slowing. This would weaken the ITC values, because the calculation of ITC is based on the across trial correlation between the phase of the oscillation of a given frequency.

Second, typing involves many responses which overlap in time. As emphasized many times before, finger presses are not executed in a serial manner in typing. If it is the initiation (rather than pressing down) of an error movement that triggers error awareness, then the phase-locking and response onset will be temporally separated. Further, the amount temporal separation (in milliseconds) is not likely to be uniform, further breaking down the correlation between the response onset and phase-locking.

These two reasons might be why we observed ITC at lower edge of the theta oscillations (i.e. 3-4Hz) rather than the higher frequencies (i.e. up to 8Hz), which are shown to undergo significant increases in ERSP. It is possible that only the oscillations which had long enough wavelengths (e.g. smaller frequencies) were found to be consistently phase locked. One complete cycle of a 4Hz oscillation lasts 250ms, which is longer than the average IKI of all participants. Compared to a 8Hz oscillation, a cycle of which lasts 125ms, phase-resetting in a 4Hz oscillation is twice as likely outlast the variability introduced by error slowing and overlapped key-presses. Even though we find this to be a plausible explanation, more empirical and computational modelling work is clearly needed to check if this speculation we offer is in fact true.

In summary, finding a statistically reliable ITC - error correction relationship was surprising and encouraging. Nevertheless, our results don't allow us to conclusively state that theta ITC is as strongly related to error correction as theta ERSP. This is because even though the corrected errors were associated with stronger theta ITC than uncorrected errors on average, this difference was not found to be statistically reliable. This may be because of the higher sensitivity of the ITC to the temporal relationship between the response locking event (pressing down of the error key) and the event that leads to the phase locking (error awareness) compared to the ERSP. Another possibility is that theta oscillations are engaged by uncorrected errors, but the timing of this engagement is not as strongly aligned with that of the error response in uncorrected errors compared to corrected errors.

5.4.3 Conclusions on Theta Oscillations and Error Correction

These results further reinforce the current opinion about the involvement of theta band oscillations in error correction and performance monitoring mechanisms. Our findings of significant theta ERSP and ITC increases in response to errors replicate previous findings (Cohen, 2011; Cohen & van Gaal, 2012; Hanslmayr et al., 2008; Luu & Tucker, 2001; Luu et al., 2003; Trujillo & Allen, 2007). Further, we show from two different angles that theta oscillations are associated strongly with error correction (rather than errors per-se). First we show that corrected errors are associated with stronger theta power and oscillations than uncorrected errors (figure 5.3). Second, we show that for the majority of participants, stronger ERSP at the time of error key-press increases the probability of error correction (figure 5.9).

Our results also support the claims of Herrojo-Ruiz et al. (2011) that error related activations can precede the onset of error instance in skilled piano players. Herrojo-Ruiz et al. (2009, 2011) suggest that internal forward models enable prediction of the sensory outcome of an error key-press and this forms the basis for detection of an error before its onset. However, we are cautious about making claims about the exact mechanism which enables such early error changes. One problem with the internal forward models assumption is that it is not clear to us how and/or when the internal forward models trigger error detection (i.e. as soon as the error response is programmed as part of an otherwise error-free chunk? As soon as the error-chunk is initiated? etc.). Second, research suggests that chunks of key-presses in skilled typing are not ballistic in nature, and typing performance can be stopped by external cues before a subsequent key is pressed (G. Logan, 1982; Rabbitt, 1978). Thus, if an error can be detected pre-initiation, the question why typists and pianists still go ahead and execute an error which is not initiated is left unanswered.

We propose on the other hand that pre-error effects on behaviour and EEG can be triggered at or after response initiation for two reasons outlined below. These points are also important points in explaining the contrasting error effects on discrete (i.e. error speeding) vs. skilled actions (i.e. error slowing). First, skilled actions such as piano playing and typing involve parallel preparation of multiple (or chunks of) key-presses. As a key is played, the fingers which will execute the next several key-presses are already moving towards their target keys (Flanders & Soechting, 1992). Thus, the time period between the initiation of key-presses and their execution in skilled actions may be long. Second, error key-presses have been shown have an increased latency compared to correct key-presses (i.e. error slowing; results from chapter 3, Herrojo-Ruiz et al. (2009); Rabbitt (1978)), extending the time elapsed between the initiation and execution of a key-press even more. Thus we believe that it is possible that the initiation of an error leads to error detection, which in turn leads both to EEG and behavioural effects before the execution of the error response.

Theta ERSP and Post-Error Behavioural Adjustment We found no relationship between the onset or the magnitude of theta ERSP peak during the error key-presses and the amount of error or post-error slowing. This is in contrast to the relationship observed between peak theta ERSP amplitude and error correction rate. As mentioned earlier in chapter 4, we see error correction as a more valid index of post-error behavioural adjustment than error slowing in the context of typing. The fact that ERSP is related to error correction but not to error slowing leads us to conclude that the ERSP is an index of performance monitoring processes timely engagement of which leads to improvements in overall performance.

Conclusions Overall, our analyses of the theta oscillations suggest that there is a reliable relationship between adaptive performance adjustments and theta parameters such as theta power and phase resetting. To reiterate, we believe theta oscillations are not necessarily sensitive to error responses, but to the subjective experience of errors. This error experience is found to be followed by error and post-error slowing. However, due to a lack of correlation between slowing measures and theta measures we are cautious about inferring a direct link between them. In short, we show here that in the skilled and continuous action of typing, theta power, and to a lesser extent theta synchronization, are reliable indices of error correction, or in more generic terms, performance monitoring.

Chapter 6

General Discussion and Conclusions

6.1 Introduction

In this final chapter we briefly summarize our findings from all chapters and provide a coherent framework within which to interpret our results as a whole. We do this by first discussing chapter-specific findings in light of findings from subsequent chapters. For example, we re-evaluate certain claims we put forward in interpreting behavioural results from the first experiment in light of insights provided by the EEG experiment. Second, we discuss our findings in relation to the points raised in the introduction chapter and with respect to current literature. Finally, we present an account to show a coherent 'big-picture' that we believe our findings show when considered together.

6.2 General Summary of Behavioural Findings

Below we summarize our findings related to the behavioural changes associated with errors. Then, we discuss certain explanations we offered for a number of patterns observed in the results presented in individual chapters. Some of these chapter specific interpretations have been supported and some dis-confirmed by findings reported in other chapters.

Note on Terminology: Error Detection vs. Error Awareness

Before moving on, it is important that we clarify two of the terms we use in the following sections. As we emphasize below, we believe that the subjective experience of an error response varies on a continuous scale. Some errors are completely missed, some errors are definitely detected and corrected, but the awareness associated with many error responses lies between these two ends of the continuum. This is probably true of many correct responses too.

We use the term error detection to refer to the experience of errors which were experienced/detected at a high subjective level which lead to error correction. In contrast, we use the term error awareness to refer to the awareness of errors which was strong enough to cause slowing of the performance (or 'orienting' (Notebaert et al., 2009)), but not strong enough to cause error correction.

We felt the need to make this distinction because we observed that corrected errors lead to error slowing as well as theta ERSP bursts, whereas uncorrected errors lead only to slowing of performance. In order for the speed of the same finger movement (matched for letter and speed) to systematically change with its accuracy, the nervous system needs to register its accuracy at some level. Thus, we believe that many uncorrected errors must have triggered some error awareness at a low level, enough to cause disturbance in typing performance and uncertainty about the accuracy of the key-press, but this awareness was not strong enough for the participant to stop typing and correct the error.

6.2.1 Pre-Error Changes in Typing Performance

In chapter 3 we reported the behavioural effects of errors on continuous typing behaviour before, during and after the error. We found that the effect of errors on typing performance was not identical in corrected and uncorrected errors. Relative to matched correct key-presses, corrected errors were slowed down and found to be preceded by faster typing behaviour. On the other hand, uncorrected error key-presses were no different than matched correct key-presses in terms of error or pre-error slowing, but were associated with strong post-error slowing.

When comparing corrected errors to uncorrected errors, the only reliable difference we found in this experiment was the amount of variability in typing speed preceding the error instance: The errors which were preceded by large variability in typing speed were less likely to be corrected than those preceded by less variability typing speed.

Based on the post-experimental comments and feedback from participants, we reasoned that the variability in typing speed might be an index of how mentally engaged they are in the typing action. As such, when participants were fully immersed in the typing action, their typing speed would be more smooth and consistent, decreasing variability in IKI; and when they are not, it would be characterized by sudden pauses, increasing variability in IKIs. Consequently, we reasoned that errors which were committed while the participant was more engaged in the task would be more likely to be detected and corrected.

However, we found no empirical support this claim in our second experiment (chapters 4 and 5): The pre-error EEG data (ERPs and brain oscillations between 2-30Hz) before corrected errors were no different than that before uncorrected errors on average. Further, in the second experiment, we failed to replicate pre-error changes in typing speed. Because of the greater statistical power of the second experiment, we conclude here that pre-error changes in performance are not as robust as error and post-error effects.

A simpler alternative explanation for the observed differences in the pre-error variability in the first experiment is related to the sample size for this analysis. Typists are more likely to detect their errors than miss them. This leads to an inherent difference in the sample sizes of corrected and uncorrected errors. Since the variance is inversely related to sample size, corrected errors are expected to be associated with smaller variance than uncorrected errors.

This difference would be exacerbated by the smaller number of trials in the first experiment as opposed to the second one. As mentioned before, the second experiment involved typing of 100 sentences as opposed to 20 in the first experiment. Inflated variance in uncorrected errors due to high error correction rate, amplified by the smaller number of data points overall in the first experiment, is likely to have contributed to the significant false positive pre-error variance effects in the first experiment.

Thus, we conclude that the increased speed and decreased variability in typing behaviour as reported in chapter 3 is more likely to be caused by the small number of observations than a robust error related behavioural effect.

6.2.2 Error Effects on Typing Performance

We found in both experiments that on average, error key-presses were slowed down compared to matched correct key-presses. In the first experiment, we found that error slowing associated with corrected errors was statistically reliable, while that associated with uncorrected errors was not. Our interpretation of this finding was that error slowing is an index of error awareness (Rabbitt, 1978). Average increase in the IKI of uncorrected errors is likely caused by the fact that at least some of the errors triggered error awareness or were unsure errors (Hewig et al., 2011).

The second experiment supported our interpretation with slight adjustments. First, error slowing for corrected errors was replicated in the second experiment. Second, error slowing for uncorrected errors was found to be smaller but statistically reliable when compared to matched correct key-presses. This pattern of results suggests that a considerable number of uncorrected errors triggered error awareness at some level; some before (small but reliable uncorrected error-slowing) and many after the error action has been completed (large post-uncorrected-error slowing).

Third, whereas almost all corrected errors were associated with error related EEG effects, only a half of uncorrected errors were associated with EEG effects (chapter 5). Furthermore, those uncorrected error key-presses which *were* associated with EEG effects, were the ones associated with the largest amount of slowing. Uncorrected errors which were not slowed down were found to be associated with none or weaker EEG effects. Coincidentally, EEG effects associated with uncorrected errors which were not slowed down were present only *after* the error key-press. This may be why post-error slowing is a much stronger effect than error slowing in uncorrected errors: Error slowing will be observed only when error awareness precedes the error commission, whereas post-error slowing can be observed even if error awareness follows the error commission.

6.3 General Summary of EEG Findings

6.3.1 Error Effects on Typing Related ERPs

One of the most important conclusions from our EEG analyses was that the relationship between ERN and error detection is stronger than is currently held in the literature (e.g. Endrass et al., 2005, 2007; Hajcak et al., 2003; Nieuwenhuis et al., 2001; OConnell et al., 2007; Shalgi, Barkan, & Deouell, 2009). We found that when compared to matched correct key-presses, corrected error ERPs are associated with very robust ERN and Pe components, and uncorrected error ERPs are associated with neither. More importantly, the corrected error ERN and Pe components were found to be reliable not only when compared to matched correct key-presses, but also when compared to uncorrected key-presses. This final point is crucial in that it suggests it is not only the Pe that is related to error detection (see a review by Overbeek, Nieuwenhuis, & Ridderinkhof, 2005, and the more recent studies cited above) so is ERN.

Traditionally, Pe has been assumed to be related to error detection and found only after detected errors, whereas ERN is related to 'error processes' and found after detected as well as undetected errors. Our results show that both ERN and Pe are associated not with error commission, but with error detection, supporting previous findings of Roger et al. (2010). It is important to emphasize here that ERN and Pe seem to reflect the same neural process in our study.

One contrast between our study and many discrete trial studies of performance monitoring in the literature is the observation of no reliable ERN in uncorrected errors. We explain why this might be the case in our discussion of current findings in relation to the literature (see section 6.4.3), and propose a framework where we link both ERN and Pe to error detection through modulation of theta power (see section 6.5).

6.3.2 Error Effects on Typing Related Theta Oscillations

Another important result was that concerning the theta oscillatory changes in response to errors and error correction. We found that corrected error key-presses were associated with stronger and reliable power bursts and synchronization in oscillations in the theta (3-8Hz) frequency range than matched correct and uncorrected error key-presses. Crucially, uncorrected errors were not followed by such strong theta power bursts. No pre-error changes were observed in the oscillations before corrected or uncorrected errors.

We believe these data indicate that theta oscillations are important indices of a performance monitoring system: When the neural processes sub-served by these oscillations become engaged (i.e become synchronized and increase in amplitude) beyond a certain threshold point, errors are likely to be corrected, improving the performance.

A number of additional analyses provided empirical support for the interpretation above. First, we found on a participant level that individual error key-presses which were associated with stronger theta power bursts were almost always more likely to be corrected than those associated with smaller theta power bursts. Second, we observed that when sorted according to the magnitude of error slowing, uncorrected error key-presses which were associated with higher theta bursts were associated with the largest magnitude of error slowing (figure 5.4). Assuming error slowing is a behavioural response to error awareness, this relationship between error slowing and theta power adds indirect empirical support to the interpretation that theta power and error awareness are linked.

6.4 Current Results With Respect to the Literature

In the introduction chapter, we summarized the error related behavioural and EEG effects and how they are reported to interact with performance monitoring. In the sections that follow, we revisit these areas as informed by our findings. Further, we include in this part directions for future research, rather than allocating a separate section to it.

6.4.1 Pre-Response Slowing and Error Detection

The observation that error key-presses are slowed down in piano playing (Herrojo-Ruiz et al., 2009) and typing (Shaffer, 1975) suggests that errors can affect continuous performance before they are completed. This is in contrast to responses in discrete trial tasks, where errors are faster than correct responses (Rabbitt, 1966b). Informed by previous findings and current results, we argue below that this contrast between error effects on discrete and continuous tasks is caused by the parallel vs. serial manner in which responses are prepared in continuous vs. discrete trial tasks.

Herrojo-Ruiz et al. (2009) and Herrojo-Ruiz et al. (2011) suggest that early error detection leads to error and pre-error slowing. Further, early error detection is likely based on internal forward models (Wolpert & Miall, 1996). Thus, error detection is triggered not by the physical initiation of the key-press, but the motor command to the effector muscles, and can influence behaviour 3 key-presses or 570ms before the error commission.

Our data supports the explanation of Herrojo-Ruiz et al. (2009) in that error key-presses were affected by error awareness before they were completed (error slowing). Our results are also compatible with the direction of causality Herrojo-Ruiz et al. (2009) proposed, in that the onset of error related EEG effects preceded the error key-press (which were the first units of behaviour to be affected by error detection - we found no pre-error slowing). However, we failed to replicate their finding of pre-error slowing and thus do not agree with their proposal that early error detection based on internal forward models causes pre-error slowing 3 key-presses before the error response.

Data Herrojo-Ruiz et al. (2009) present to claim that error detection leads to *pre-error* slowing don't support this claim neither. Herrojo-Ruiz et al. (2009) show that the onset of pre-error speeding was about 570ms before the error commission, whereas the onset of the earliest error related EEG effects was about 120ms before. It is not clear to us how the effect of error detection on behaviour (pre-error slowing, at -570ms) can precede the onset of the indices of its generation (EEG effects at -120ms).

We believe that a simpler explanation for pre-response error detection is at least as well equipped to explain how errors can be detected before they are executed, based on a number of previous findings. Below we argue (after presenting empirical support) that there is enough time between response initiation and response execution in skilled typing to allow for the slowing down of an error key-press.

First, Flanders and Soechting (1992) and Soechting and Flanders (1992) report that bi-manual key-presses in typing start to move 200ms before the time they execute the key-press. This suggests that on average, 200ms is elapsed between the initiation of a finger movement and the pressing down of the key.

Second, Fischer and Ramsperger (1984) show that 100 to 150ms is a long enough time for the sensory, perceptual and motor systems to perceive and process the location of a visually presented stimulus, and then programme an accurate saccade to its location. We acknowledge that making saccades and key-presses are likely to involve different neural pathways. However, the report of Fischer and Ramsperger (1984) is important in showing that the lower limit of time required by the neural system to process an input and implement an appropriate behavioural response to its location (in a similar manner to choice reaction time tasks in that the location of the visual stimulus is not predictable but determines the accuracy of the response) is in the 100ms - 150ms range.

Third, and most importantly, G. Logan (1982) showed that an external tone (i.e. not predictable by internal forward models) can be used to interrupt typing of single words. G. Logan (1982) found that his participants could stop typing in response to the tone within 250-300ms. This suggests that the ballistic units of typing (if any) are no larger than one or two (chunked) key-presses, and typing can be interrupted very shortly after the presentation of an external stimulus.

Findings of Flanders and Soechting (1992); Soechting and Flanders (1992) and G. Logan (1982) suggest that even though key-presses in typing are programmed in parallel (Flanders & Soechting, 1992; Soechting & Flanders, 1992), they are 'monitored' serially. The key factor in the ability to disrupt individual key-presses in time, as also implied by G. Logan (1982), is the time period between the stimulus that dictates disruption (be it an external tone or internal error detection), and the next response in the action sequence.

Based on the points raised above, we propose that error detection can be triggered by somato-sensory feedback from fingers at the time of error initiation, which in turn slows down the error key-presses. It follows that other error detection related parameters (e.g. theta power) being constant: i) Errors with IKIs faster than a low threshold will be less likely to be slowed, ii) Errors with IKIs slower than an upper threshold will be stopped/cancelled before they are executed, and iii) Those errors which are between these two limits will be subject to error slowing.

We would like to point out that we do not argue that internal forward models have no role to play in the execution or even in the monitoring of skilled actions like typing or piano playings. Internal forward models, for example, might provide important input to the neural processes which represent the IKI thresholds mentioned in the previous paragraph.

Rather, we argue that the pre-error slowing observed by Herrojo-Ruiz et al. (2009) is not caused by early error detection since 1) Error related neural activity doesn't precede pre-error slowing (Herrojo-Ruiz et al., 2011), 2) There is evidence that chunks of multiple key-presses in skilled actions are not ballistic to the extent that an error detected 3 key-presses or 570ms before its execution is slowed but cannot be cancelled, and 3) Time elapsed between the onset and the execution of key-presses in skilled actions potentially allow enough time for error detection to slow down the execution of those responses, and provide a simpler account which fits the error and post-error slowing observations accumulated over decades more closely.

6.4.2 Post-error slowing

We found in line with G. Logan and Crump (2010) that error slowing was associated with uncorrected errors as well as corrected errors. Further, we found that the strength of error related EEG effects, while predicting error correction, was not related to the amount of post-error slowing. These two findings suggest to us that post-error slowing is not an adaptive behavioural change caused by error detection. Rather we see post-error slowing as an 'attentional orientation' response (Notebaert et al., 2009), triggered by errors of which the participants are 'unsure' (Hewig et al., 2011).

Thus, our data are consistent with those of Hewig et al. (2011), who found that error detection was strongly related to the ERP measures such as the ERN and Pe, but not to post-error slowing; and also with the views of Notebaert et al. (2009) that post-error slowing is not an adaptive effect.

Post-Error Slowing as a Response to Weak Error Awareness

So far, we argued that ERN, Pe, and theta oscillations are robust indices of error detection and awareness. Also, we pointed out earlier that ERN is rarely found to be related to error detection, since it is reliably observed after undetected errors. Here we provide an explanation for this contrast between our results and those reported in the literature.

We propose here a threshold mechanism which may be involved in modulating the relationship between error detection/correction and error related EEG and behavioural effects. It is important to emphasize here that two important factors in determining which activations cross a given threshold are the *duration* and *strength* of activation (i.e. a strong burst of activity lasting 10ms is less likely to cross a given threshold than slightly weaker burst of activity lasting 50ms).

For simplicity, consider 4 scenarios in our experiment. In scenario 1, the participant made an error response and immediately corrected it. In scenario 2, the she made an error and corrected it after pressing 1 or 2 key-presses. In scenario 3, she made a mistake, but kept typing without correcting it because she was not sure if it was an error. In scenario 4, she made an error and kept typing because she didn't even notice she made an error.

Now for the sake of the argument, assume that the theta oscillations are greatest in scenario 1 (theta ERSP increase = 10dB); slightly weaker in scenario 2 (theta ERSP increase = 8dB); weak in scenario 3 (theta ERSP = 3dB); and minimal in scenario 4 (theta ERSP burst = 1dB). These scenarios are based on the current findings summarized in figures 5.1, 5.2, 5.3 and 5.4.

We propose that there two thresholds the theta ERSP response needs to cross in order to have any effect on performance. Crossing the first, lower threshold leads to a subjective state of uncertainty about whether an error has been committed. Crossing the second, higher threshold leads to error detection and correction via the backspace. Our results suggest that even a small theta ERSP response is enough to cross the first threshold, and *to slow the performance down.* However, in order for a theta response to cross the second threshold, it should be strong and sustained for a minimal amount of time. This, we speculate, is the reason why uncorrected errors are associated with reliable error and post-error slowing in the absence of reliable error related ERP or oscillatory effects. A question that remains then, is: Why do undetected errors in discrete trials lead to an ERN and those in typing don't?

Absence of ERN in Uncorrected Typing Errors

Within the above frame-work, corrected typing errors would be those that trigger i) Strong perturbation of theta power (scenario 1), and ii) Slightly weaker, but longer lasting perturbation of theta power (scenario 2). Uncorrected errors on the other hand would be those that trigger i) Weak and short theta ERSP (scenario 3) and ii) No theta ERSP (scenario 4). However, in a discrete trial task where errors need to be signalled verbally (or via another button press), the second threshold might be higher. This is because the association between the error awareness and pressing of backspace, strengthened over 100s of hours of typing experience is not present in experimental tasks such as the flankers or go/no-go. One can test the validity of this claim by trying to signal their typing errors verbally (a non-trained error signalling response) vs. by pressing the backspace (a highly trained error signalling response).

This more conservative threshold for error signalling would mean that the uncorrected errors in discrete trial tasks would contain a greater proportion of scenario 2 errors, with short latency / sub-threshold bursts of theta ERSP. In other words, many uncorrected errors would be associated with stronger but not sustained theta ERSP increases. When averaged together, the earlier parts of the oscillatory activity would be expressed in the grand average ERPs (e.g. the first half cycle), whereas the later parts (e.g. the second half cycle) would not because the theta perturbation was not sustained. This would lead to a smaller but reliable ERN, and no Pe, in the average ERP of undetected errors in discrete trial tasks.

One important prediction of the threshold mechanism presented above is that if the threshold for error correction/signalling is lower in typing, then more of the scenario 2 errors should be included in the corrected errors in our experiment than in discrete trial tasks. Accordingly, this should be evident in the corrected error ERP record. If the corrected errors contain errors which lead to weaker and longer lasting theta ERSP bursts, than the ERP record should reflect more cycles of the theta oscillation than discrete trial tasks. Figures 4.8 and 4.9 show that this may be the case particularly when compared to the typical ERN resulting from discrete trial tasks (figure 1.1).

Future Direction

We acknowledge that the lack of relationship between post-error slowing and theta measures that we use in support of our claim that these two measures are not causally related are based on a null effect (no significant correlation between post-error slowing and theta power across trials). Future experiments can be designed to examine whether oddball events other than errors will have the same effect on post-slowing and errors when matched for frequency (and other factors found to effect the orientating response). A study of patients with frontal lobe damage by Gehring and Knight (2000), which revealed that post-error slowing and adaptive changes to post-error performance dissociate, suggests that the results of such experiments can be promising.

With respect to the threshold based error detection mechanism proposed above, we cannot exclude the fact that the uncorrected errors had a smaller sample size than corrected errors. This relative difference in statistical power might be why we observed no ERN in uncorrected errors. An experiment with a more balanced number of corrected and uncorrected error instances might be sufficient to falsify the hypothesis proposed above by presenting a statistically reliable uncorrected-ERN. On the other hand, computational modelling might be a key-approach to test more predictions of the mechanism proposed above in an attempt to provide support for its validity.

6.4.3 Error awareness, Pe, Theta Oscillations

Of all the error related parameters we used, we found that theta oscillatory power was the one most strongly linked to adaptive changes in performance (i.e. error correction). In the introduction chapter section 1.2.3, we stated that the current literature suggested that the Pe amplitude was also related to error detection, with a correlation between the Pe amplitude and error detection.

We would like to put forth a speculation linking Pe and theta power, even though we present no modelling work to support it. We have shown that almost all corrected error responses lead to strong and sustained increases in theta power, which lasts for more than half a second (from 98ms before 465ms after the response at 4Hz oscillations, results from chapter 5). This is in contrast to the much weaker association between theta power and uncorrected errors. So, we believe the response of theta power to a given error is predictive of the correction of that error, such that the stronger the theta, the higher the likelihood of error correction.

Strong theta power bursts which last until almost half a second after the error response are likely to cause substantial changes in the ERPs recorded during that time. In table 6.1 we present the onset and duration of theta and ERP effects of errors. As a fore-warning, we would like to point out that the temporal resolution of the output from the wavelet transform (i.e. ERSP onset) is not as precise as that of the ERPs (i.e. ERN and Pe), particularly for oscillations of lower frequencies (Herrmann et al., 2005). Thus it is important to remember that the margin of error in ERSP onset is expected to be slightly greater than that of ERPs onsets while interpreting the information presented in table 6.1.

EEG Measure	Onset (ms)	Offset (ms)	Duration (ms)	Frequency
ERN	-156	20	176	6Hz
Pe	70	211	141	7Hz
ERN + Pe	-156	211	367	3Hz
ERSP @ 3Hz	-79	449	528	
ERSP @ 4Hz	-98	456	554	
ERSP @ 6Hz	-148	332	480	
ERSP @ 8Hz	-129	566	695	

Table 6.1: Table showing the onset and offset of EEG effects relative to the time of corrected error response, and for the ERP effects, the frequency their duration corresponds to.

The data provided in table 6.1 show that, 1) the onset of ERP and theta effects largely overlap and 2) the oscillatory effects outlast the ERP effects (but see the next paragraph). Further, as the durations of ERN and Pe are comparable it is possible that they are two half-cycle manifestations of the same oscillatory process, the wavelength of which is 367ms (3Hz).

Even though the oscillatory effects seem to outlast ERP effects, a closer look at figure 4.8 shows that the 3Hz oscillatory pattern evident in the corrected error ERP lasts well into 500ms post-response where it starts to deteriorate. As presented in figure 4.9, there were time points within the 350-550ms post-response period which were found to be temporarily reliable after correction for multiple comparisons.

Thus we would like to entertain the possibility that errors which lead to phase-locked increases in power at these oscillations, lead to error detection and subsequently to error correction. Those which do not lead to such strong oscillatory changes only lead to slowing without any immediate adaptive changes in the behaviour. The ERN and Pe are the ERP manifestations of increases in power and synchrony in these oscillations.

We conclude chapter by presenting and argument to encapsulate the important findings from different chapters as a consistent whole in the next section.

6.5 Conclusions - The Big Picture

Here, we present a coherent account reflecting our interpretation of the results reported in this thesis, as informed by the literature on the relationship between theta oscillations and performance monitoring and error detection (Cavanagh et al., 2009, 2011; Cohen, 2011; Cohen & van Gaal, 2012; Cohen et al., 2009). Our aim here is to provide a framework within which the significant behavioural and EEG results compliment each other and when considered together, informs our current understanding of how the detection of typing errors evolves over time and when they lead to adaptive changes in performance and when they do not.

We propose that lower theta (3-4Hz) neural oscillations constitute a crucial component of performance monitoring ability. In the event of a change in the environment (such as an error) which necessitates changes to on-going behaviour, these lower theta oscillations 'fire', empirically quantified as synchronized (ITC) bursts of spectral power (ERSP).

If this 'firing' crosses a threshold level of magnitude and is sustained, the online behaviour is interrupted and necessary behavioural changes are made (strong ERSP and ITC in corrected errors). If the 'firing' is not strong enough (smaller or no ERN or ERSP during uncorrected errors), or is not sustained (no Pe after undetected errors in discrete response tasks, shorter duration of ERSP burst in uncorrected errors, figure 5.4), then the errors are not followed by adaptive changes in ongoing behaviour.

Sub-threshold or post-response 'firing' might be enough to slow the performance down and cause an attentional orientation response (e.g., uncorrected-error-slowing, (Notebaert et al., 2009)), but unless it is sustained for a long enough time to engage the higher level cognitive control processes, ongoing behaviour may not be adjusted in an adaptive manner.

Future Directions

If it is true that the neural process sub-serving performance monitoring and error detection mechanisms are oscillations at 3-4Hz frequency centralized over the ACC, the disruption of these oscillations should disrupt adaptive performance changes. For example, the disruption of these oscillations over the ACC (as opposed to other frequencies over ACC), using repetitive trans-cranial stimulation (rTMS), should hypothetically affect error detection rates while leaving the error rate unchanged. It is possible that rTMS at these oscillations on sites other than the ACC such as the lateral pre-frontal cortex (Gehring & Knight, 2000) might affect error correction as well. Such findings would corroborate EEG findings that show the importance of functional connectivity and interaction between the medial and lateral areas in modulating error related changes (e.g., Cavanagh et al., 2009; Cohen, 2011; Cohen & van Gaal, 2012; Cohen et al., 2009).

6.6 General Conclusions

In the beginning of this thesis, we pointed out to the importance of performance monitoring and the benefits of a better understanding of its neural and behavioural markers. We asserted that by studying how these markers interact we may start to understand how and when errors are processed and more importantly, adaptive adjustments to on-going performance are implemented. In this chapter, we provided a coherent account of how typing errors may cause changes in the typing speed, what electro-physiological markers are sensitive these to errors, and how the response of these electro-physiological markers are predictive of the behavioural response of the typist to the error.

As pointed out in the future directions sections, more research is definitely needed to provide more support for (or falsification of) our findings. Nevertheless, we show here with confidence that behavioural and EEG data can be acquired in ecological tasks like typing in the study of performance monitoring, and reliable results can be acquired with careful design and analysis. We believe this was the most important outcome of the current work: Use of ecological tasks in psychological research holds great potential for informing our understanding of how our nervous system controls and adapts our actions online.

References

Alexander, W. H., & Brown, J. W. (2010). Computational models of performance monitoring and cognitive control. *Topics in cognitive science*, *2*(4), 658–677.

Allain, S., Carbonnell, L., Falkenstein, M., Burle, B., & Vidal, F. (2004). The modulation of the ne-like wave on correct responses foreshadows errors. *Neuroscience letters*, *372*(1-2), 161-166.

Armbrecht, A. S., Wohrmann, A., Gibbons, H., & Stahl, J. (2010). Temporal dynamics of conflict monitoring and the effects of one or two conflict sources on error-related negativity. *Biological Psychology*, *85*(1), 112–123.

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., , & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (mri) and functional mri. *Journal of Neuroscience*, *27*, 3743 – 3752.

Botvinick, M., Braver, T., Barch, D., Carter, C., & Cohen, J. (2001). Conflict monitoring and cognitive control. *Psychological review*, *108*(3), 624.

Burle, B., Roger, C., Allain, S., Vidal, F., & Hasbroucq, T. (2008). Error negativity does not reflect conflict: A reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience*, *20*(9), 1637-1655.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*(6), 215–222.

Buzsaki, G. (2006). *Rhythms of the brain.* Oxford University Press.

Campagne, A., Pebayle, T., & Muzet, A. (2004). Correlation between driving errors and vigilance level: influence of the drivers age. *Physiology and Behaviour*, *80*, 515–524.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*, 747–749.

Castellar, E., Kuhn, S., Fias, W., & Notebaert, W. (2010). Outcome expectancy and not accuracy determines posterror slowing: Erp support. *Cognitive Affective and Behavioral Neuroscience*, *10*, 270-278.

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. B. (2009). Prelude to and resolution of an error: Eeg phase synchrony reveals cognitive control dynamics during action monitoring. *The Journal of Neuroscience*, *29*(1), 98–105.

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioural adaptation in reinforcement learning. *NeuroImage*, *49*(4), 3198–3209.

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2011). Theta lingua franca: A common mid-frontal substrate for action monitoring processe. *Psychophysiology*, *49*(2), 220-238.

Chevrier, A., & Schachar, R. J. (2010). Error detection in the stop signal task. *Neuroimage*, *53*, 664 – 673.

Cohen, M. X. (2011). Error-related medial frontal theta activity predicts cigulate-related structural connectivity. *Neuroimage*, *55*, 1373–1383.

Cohen, M. X., & van Gaal, S. (2012). Dynamic interactions between large-scale brain networks predict behavioral adaptation after perceptual errors. *Cerebral Cortex*.

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R. K., & Lamme, V. A. F. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Frontiers in human neuroscience*, *3*(54).

Coles, G. H., Michael, Scheffers, M. K., & Fournier, L. (1995). Where did you go wrong? errors, partial errors, and the nature of human information processing. *Acta Psychologica*, *90*, 129 – 144.

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, *2*, 233.

de Bruijn, E. R. A., Hulstijn, W., Meulenbroek, R. G. J., & van Galen, G. P. (2003). Action monitoring in motor control: Erps following selection and execution errors in a force production task. *Psychophysiology*, *40*(5), 786-95.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *The Journal of Neuroscience*, *25*(50), 11730-11737.

Delorme, A., & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics. *Journal of Neuroscience Methods*, *134*, 9 – 21.

Desmet, C., Imbo, I., de Brauwer, J., Brass, M., Fias, W., & Notebaert, W. (2012). Error adaptation in mental arithmetic. *The Quarterly Journal of Experimental Psychology*, *65*(6), 1059-1067.

Ehlis, A.-C., Herrmann, M. J., Bernhard, A., & Fallgatter, A. J. (2005). Monitoring of internal and external error signals. *Journal of Psychophysiology*, *19*(4), 263-269.

Eichele, H., Juvodden, H. T., Ullsperger, M., & Eichele, T. (2010). Mal-adaptation of event-related eeg responses preceding performance errors. *Frontiers in human neuroscience*, *4*(1), 1–9.

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, K, A., … Ullsperger, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16), 6173–6178.

Endrass, T., Franke, C., & Kathmann, N. (2005). Error awareness in a saccade countermanding task. *Journal of Psychophysiology*, *19*(4), 275-280.

Endrass, T., Reuter, B., & Kathmann, N. (2007). Erp correlates of conscious error recognition: aware and unaware errors in an antisaccade task. *European Journal of Neuroscience*, *26*, 1714-1720.

Ericsson, K. A., & Krampe, R. T. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363–406.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance:

Evidence on maximal adaptations on task constraints. *Annual Review of Psychology*, *47*, 273 – 305.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the indentification of a target letter in a nonsearch task. *Perception and Psychophysics*, *16*(1), 143–149.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late erp components. ii. error processing in choice reaction tasks. *Electroencephalography and clinical neurophysiology*, *78*(6), 447–455.

Fiehler, K., Ullsperger, M., & von Cramon, D. Y. (2004). Neural correlates of error detection and error correction: is there a common neuroanatomical substrate? *European Journal of Neuroscience*, *19*, 3081–3087.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, *57*, 191 - 195.

Flanders, M., & Soechting, J. F. (1992). Kinematics of typing: parallel control of the two hands. *Journal of neurophysiology*, *67*(5), 1264–1274.

Fournier, L. R., Scheffers, K., Coles, G. H., Michael, Adamson, A., & Abad, V., Enrique. (1997). The dimensionality of the flanker compatibility effect: A psychophysiological analysis. *Psychology Research*, *60*, 144 – 155.

Gehring, W. J., & Fencsik, D. E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *The Journal of Neuroscience*, *21*(23), 9430–9437.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychologcial Science*, *4*(6), 385-390.

Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Slamming on the brakes: An electrophysiological study of error response inhibition. *Psychological Science*, *11*, 1 – 6.

Gehring, W. J., & Knight, R. T. (2000). Prefrontalcingulate interactions in action monitoring. *Nature Neuroscience*, *3*(5), 516–520.

Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The error related negativigy..

Gentner, D. R. (1983). Keystroke timing in transcription typing. In W. E. Cooper (Ed.), *Cognitive aspects of skilled typewriting* (pp. 95–117). New York: Springer-Verlag.

Gentner, D. R. (1984, April). *Expertise in typewriting.* (Tech. Rep. No. ADA142865). California University San Diego La Jolla Center For Human Information Processing.

Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self-and externally caused failure of goal achievement. *Neuroimage*, *47*(4), 2023–2030.

Gerloff, C., Corwell, B., Chen, R., Hallett, M., & Cohen, L. G. (1997). Stimulation over the human supplementary motor area interferes with the organization of future elements in complex motor sequences. *Brain*, *120*(9), 1587-1602.

Groppe, D. M., Makeig, S., & Kutas, M. (2008). Independent component analysis of event-related potentials. *Cognitive Science Online*, *6*(1), 1–44.

Grundler, T. O. J., Cavanagh, J. F., Figueroa, C. M., Frank, M. J., & Allen, J. J. B. (2009). Task-related dissociation in ern amplitude as a function of obsessive-compulsive symptoms. *Neuropsychologia*, *47*(8–9), 1978–1987.

Hajcak, G., Mcdonald, N., & Simons, R. F. (2003). To err is autonomic: Error-related brain potentials, ans activity, and post-error compensatory behavior. *Psychophysiology,*, *40*, 895–903.

Hajcak, G., Nieuwenhuis, S. T., Ridderinkhof, R. K., & Simons, F., Robert. (2005). Error-preceding brain activity: Robustness, temporal dynamics, and boundary conditions. *Biological Psychology*, *70*, 67–78.

Hajcak, G., & Simons, R. F. (2008). Oops!.. i did it again: An erp and behavioral study of double-errors. *Brain and Cognition*, *68*(1), 15–21.

Hanslmayr, S., Pastotter, B., Bauml, K.-H., Gruber, S., Wimber, M., & Klimesch, W. (2008). The electrophysiological dynamics of interference during the stroop task. *Journal of Cognitive Neuroscience*, *20*(2), 215-225.

Heldman, M., Russeler, J., & Munte, T. F. (2008). Internal and external information in error processing. *BMC Neuroscience*, *9*(1), 33.

Herrmann, C. S., Grigutsch, M., & Busch, N. A. (2005). Eeg oscillations and wavelet analysis. MIT Press. Retrieved from `http://edoc.mpg.de/261589`

Herrojo-Ruiz, M., Jabusch, H.-C., & Altenmuller, E. (2009). Detecting wrong notes in advance: neuronal correlates of error monitoring in pianists. *Cerebral cortex*, *19*(11), 2625–2639.

Herrojo-Ruiz, M., Strbing, F., Jabusch, H.-C., & Altenmuller, E. (2011). Eeg oscillatory patterns are associated with error prediction during music performance and are altered in musician's dystonia. *NeuroImage*, *55*, 1791–1803.

Hester, R., Foxe, J. J., Molholm, S., Shpaner, M., & Garavana, H. (2005). Neural mechanisms involved in error processing: A comparison of errors made with and without awareness. *NeuroImage*, *27*, 602–608.

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

Hoffmann, S., & Falkenstein, M. (2010). Independent component analysis of erroneous and correct responses suggests online response control. *Human Brain Mapping*, *31*, 1305–1315.

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

Holroyd, C. B., Dien, J., & Coles, M. G. H. (1998). Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neuroscience letters*, *242*(2), 65–68.

Howell, D. C. (2002). Resampling and nonparametric approaches to data. In *Statistical methods for psychology* (chap. 18). Duxbury.

Huster, R. J., Eichele, T., Enriquez-Geppert, S., Wollbrink, A., Kugel, H., Konrad, C., & Pantev, C. (2011). Multimodal imaging of functional networks and event-related potentials in performance monitoring. *NeuroImage*, *56*, 1588–

1597.

Inhoff, A. W. (1991). Word frequency during copytyping. *Journal or Experimental Psychology: Human Perception and Performance*, *17*(2), 478 – 487.

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*, 1395–1399.

Kayser, J., & Tenke, C. (2006). Principal components analysis of laplacian waveforms as a generic method for identifying erp generator patterns: I. evaluation with auditory oddball tasks. *Clinical Neurophysiology*, *117*, 348-368.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. *Technical Training*, *Research Branch Report*(February), 8–75.

Kuhn, A. A., Williams, D., Kupsch, A., Limousin, P., Hariz, M., Schneider, G. H., ... Brown, P. (2004). Event-related beta desynchronization in human subthalamic nucleus correlates with motor performance. *Brain*, *127*, 735 – 746.

Laming, D. (1979). Choice reaction performance following an error. *Acta Psychologica*, *43*(3), 199–224.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (p. 112-136). New York: Wiley.

Laukka, S. J., Jarvilehto, Alexandrov, Y. I., & Lindqvist, J. (1995). Frontal midline theta related to learning in a simulated driving task. *Biological Psychology*, *40*, 313–320.

Lee, T. W., Girolami, M., & Sejnowski, T. (1999). Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and super-gaussian sources. *Neural Computation*, *11*(2), 609 – 633.

Leuthold, H., & Sommer, W. (1999). Erp correlates of error processing in spatial s-r compatibility tasks. *Clinical Neurophysiology*, *110*(2), 342-357.

Logan, F. A. (1999). Errors in copy typewriting. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1760–1773.

Logan, G. (1982). On the ability to inhibit complex movements: A stop-signal study of typewriting. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(6), 778–792.

Logan, G., & Crump, M. (2009). The left hand doesn't know what the right hand is doing: the disruptive effects of attention to the hands in skilled typewriting. *Psychological Science*, *20*(10), 1296–300.

Logan, G., & Crump, M. (2010). Cognitive illusions of authorship reveal hierarchical error detection in skilled typists. *Science*, *330*(6004), 683–686.

Logan, G., & Crump, M. J. C. (2011). Hierarchical control of cognition and action: The case for skilled typewriting. In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 54, pp. 1–27). Academic Press.

Luu, P., & Tucker, D. M. (2001). Regulating action: alternating activation of midline frontal and motor cortical networks. *Clinical Neurophysiology*, *112*, 1295–1306.

Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychological Science*, *14*(1), 47–53.

Luu, P., Tucker, D. M., & Makeig, S. (2004). Frontal midline theta and the error-related negativity: neurophysiological mechanisms of action regulation. *Clinical Neurophysiology*, *115*, 1821–1835.

MacNeilage, P. F. (1964). Typing errors as clues to serial ordering mechanisms in language behaviour. *Language and Speech*, *7*, 144 – 159.

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: Erp effects prior to performance errors in musicians indicate fast monitoring processes. *PLoS ONE*, *4*(4), 5032.

Maier, M. E., Steinhauser, M., & Hubner, R. (2010). Effects of response-set size on error-related brain activity. *Experimental Brain Research*, *202*, 571–581.

Maier, M. E., Yeung, N., & Steinhauser, M. (2011). Error-related brain activity and adjustments of selective attention following errors. *NeuroImage*, *56*(4), 2339–2347.

Makeig, S., Jung, T. P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences of the U. S. A.*, *94*, 10979-10984.

Marco-Pallares, J., Camara, E., Munte, T. F., & Rodriguez-Fornells, A. (2008). Neural mechanisms underlying adaptive actions after slips. *Journal of Cognitive Neuroscience*, *20*, 1595 – 1610.

Mazaheri, A., Nieuwenhuis, I. L. C., van Dijk, H., & Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Human Brain Mapping*, *30*, 1791–1800.

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

Miltner, W. H. R., Lemke, U., Weiss, T., Holroyd, C., Scheffers, M. K., & Coles, M. G. (2003). Implementation of error-processing in the human anterior cingulate cortex: a source analysis of the magnetic equivalent of the error-related negativity. *Biological Psychology*, *64*, 157-166.

Mizuki, Y., Tanaka, M., Isozaki, H., Nishijima, H., & Inanaga, K. (1980). Periodic appearance of theta rhythm in the frontal midline area during performance of a mental task. *Electroencephalography and Clinical Neurophysiology*, *49*, 345–351.

Murphy, P. R., Robertson, I. H., Allen, D., Hester, R., & O'Connell, R. G. (2012). An electrophysiological signal that precisely tracks the emergence of error awareness. *Frontiers in human neuroscience*, *6*(65).

Neubert, F. X., Mars, R. B., Buch, E. R., Olivier, E., & Rushworth, M. F. (2010). Cortical and subcortical interactions during action reprogramming and their related white matter pathways. *Proceedings of National Academy of Science U.S.A.*, *107*, 13240 – 13245.

Nieuwenhuis, S. T., Ridderinkhof, K., Richard, Talsma, D., Coles, M. G., Holroyd,

C. B., Kok, A., & Van Der Molen, M. (2002). A computational account of altered error processing in older age: Dopamine and error-related negativity. *Cognitive, Affective and Behavioural Neuroscience*, *2*(1), 19–36.

Nieuwenhuis, S. T., Ridderinkhof, R. K., 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*, 752–760.

Norman, D., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. S. R Davidson & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4). Plenum Press.

Notebaert, W., Houtman, F., van Opstal, F., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, *111*, 275-279.

Olvet, D., & Hajcak, G. (2009). The stability of error-related brain activity with increasing trials. *Psychophysiology*, *46*, 1–5.

Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. In Neuper & Klimesch (Eds.), *Progress in brain research* (Vol. 159). Elsevier.

Oostendorp, T. F., & van Oosterom, A. (1989). Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. *IEEE Transactions on Biomedical Engineering*, *36*(3), 382-391.

Overbeek, T. J. M., Nieuwenhuis, S. T., & Ridderinkhof, R. K. (2005). Dissociable components of error processing: On the functional significance of the pe vis--vis the ern/ne. *Journal of Psychophysiology*, *19*, 319-329.

OConnell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H., ... Foxe, J. J. (2007). The role of cingulate cortex in the detection of errors with and without awareness: a high-density electrical mapping study. *European Journal of Neuroscience*, *25*, 2571-2579.

Pailing, P. E., Segalowitz, S. J., Dywan, J., & Davies, P. L. (2002). Error negativity and response control. *Psychophysiology*, *39*, 198–206.

Rabbitt, P. (1966a). Error correction time without external error signal. *Nature*, *212*(22), 438–438.

Rabbitt, P. (1966b). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, *71*(2), 264–272.

Rabbitt, P. (1967). Time to detect errors as a function of factors affecting choice-response time. *Acta Psychologica*, *27*, 131–142.

Rabbitt, P. (1968). Three kinds of error-signalling responses in a serial choice task. *The Quarterly Journal of Experimental Psychology*, *20*(2), 179–188.

Rabbitt, P. (1978). Detection of errors by skilled typists. *Ergonomics*, *21*(11), 945–958.

Rabbitt, P., & Rodgers, B. (1977). What does a man do after he makes an error? an analysis of response programming. *The Quarterly Journal of Experimental*, *29*(4), 727–743.

Reason, J. (1990). *Human error.* Cambridge University Press.

Reason, J. (2000). Human error: models and management. *British Medical Journal*, *320*, 768 – 770.

Ridderinkhof, K., Richard. (2002). Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Pyschological Research*, *66*, 312 – 323.

Ridderinkhof, R. K., Nieuwenhuis, S. T., & Bashore, T. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neuroscience letters*, *348*(1), 1–4.

Rodriguez-Fornells, A., Kurzbuch, A. R., Munte, T. F., & Munte, T. F. (2002). Time course of error detection and correction in humans: neurophysiological evidence. *The Journal of Neuroscience*, *22*(22), 9990–9996.

Roger, C., Benar, C. G., Vidal, F., Hasbroucq, T., & Burle, B. (2010). Rostral cingulate zone and correct response monitoring: Ica and source localization evidences for the unicity of correct- and error-negativities. *Neuroimage*, *51*(1), 391–403.

Rosenbaum, D. A. (1991). *Human motor control.* Academic Press.

Sadato, N., Yonekura, Y., Waki, A., Yamada, H., & Ishii, Y. (1997). Role of the supplementary motor area and the right premotor cortex in the coordination of bimanual finger movements. *The Journal of Neuroscience*, *17*(24), 9667-9674.

Salthouse, A., Timothy. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. *Psychological Bulletin*, *99*(3), 303 – 319.

Scheffers, M. K., Coles, M. G. H., Bernstein, P., Gehring, W. J., & Donchin, E. (1996). Event-related brain potentials and error-related processing: An analysis of incorrect responses to go and not-go stimuli. *Psychophysiology*, *33*(1–2), 42–53.

Shaffer, L. H. (1975). Control processes in typing. *The Quarterly Journal of Experimental Psychology*, *27*(3), 419–432.

Shaffer, L. H., & Hardwick, J. (1969). Reading and typing. *The Quarterly Journal of Experimental Psychology*, *21*(4), 381-383.

Shalgi, S., Barkan, I., & Deouell, L. Y. (2009). On the positive side of error processing: error-awareness positivity revisited. *European Journal of Neuroscience*, *29*(7), 1522–1532. Retrieved from `http://dx.doi.org/10.1111/j.1460-9568.2009.06690.x` doi: 10.1111/j.1460-9568.2009.06690.x

Skinner, B. F. (1959). *Cumulative record.* New York : Appleton-Century-Crofts.

Smith, G. A., & Brewer, N. (1995). Slowness and age: Speed-accuracy mechnisms. *Psychology and Aging*, *10*(2), 238–247.

Soechting, J. F., & Flanders, M. (1992). Organization of sequential typing movements. *Journal of Neurophysiology*, *67*, 1275 - 1290.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662.

Swann, N., Tandon, N., Canolty, R., Ellmore, T. M., Mcevoy, L. K., Dreyer, S., ... Aron, A. R. (2009). Intracranial eeg reveals a time- and frequency-specific role for the right inferior frontal gyrus and primary motor cortex in stopping initiated responses. *Journal of Neuroscience*, *29*, 12675 – 12685.

Trujillo, L. T., & Allen, J. B., John. (2007). Theta eeg dynamics of the error-related negativity. *Clinical Neurophysiology*, *118*, 645–668.

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*, 1387–1401.

Ullsperger, M., & von Cramon, D. Y. (2006). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fmri and erps. *Brain research*, *5*(1), 102–109.

Ursu, S., Clark, K. A., Aizenstein, H. J., Stenger, V. A., & Carter, C. S. (2009). Conflict-related activity in the caudal anterior cingulate cortex in the absence of awareness. *Biological Psychology*, *80*, 279-286.

van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of cognitive neuroscience*, *14*(4), 593–602.

van Ede, F., De Lange, F., Jensen, O., & Maris, E. (2011). Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha- and beta-band oscillations. *Journal of Neuroscience*, *31*, 2016 – 2024.

Van't Ent, D., & Apkarian, P. (1999). Motoric response inhibition in nger movement and saccadic eye movement: a comparative study. *Clnical Neurophysiology*, *110*, 1058–1072.

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *NeuroImage*, *14*, 1302-1308.

Vidal, F., Burle, B., Bonnet, M., Grapperon, J., & Hasbroucq, T. (2003). Error negativity on correct trials: a reexamination of available data. *Biological Psychology*, *64*(3), 265–282.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity' specific to errors? *Biological psychology*, *51*(1–2), 109–128.

Wells, F. L. (1916). On the psychomotor mechanisms of typewriting. *The American Journal of Psychology*, *27*(1), 47–70.

Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence. *Frontiers in Human Neuroscience*, *6*(88).

Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and error: Common neuronal architecture for the processing of errors and novelty. *The Journal of Neuroscience*, *32*(22), 7528–7537.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67–85.

Williams, M. A., & Ericsson, K., Anders. (2005). Perceptual-cognitive expertise in sport: Some considerations when applying the expert performance approach. *Human Movement Science*, *24*(3), 283 – 307.

Willoughby, A. (2012). *galvani's frog.* `http://galvanisfrog.com/Research.php`. ([Online; accessed 1-July-2012])

Wolpert, D. M., & Miall, R. C. (1996). Forward models for physiological motor control. *Neural Networks*, *9*(8), 1265–1279.

Yeung, N., Bogacz, R., Holroyd, C. B., Nieuwenhuis, S. T., & Cohen, J. D. (2004). Theta phase resetting and the error-related negativity. *Psychophysiology*, *44*, 39–49.