MEDIAL FRONTAL BRAIN POTENTIALS FOLLOWING FEEDBACK DURING PROBABILISTIC LEARNING

by

Adrian R. Willoughby

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in The University of Michigan
2005

Doctoral Committee:

Associate Professor William J. Gehring, Chair
Professor David E. Meyer
Professor Patricia A. Reuter-Lorenz
Associate Professor Stephan F. Taylor

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 3163969

Copyright 2005 by
Willoughby, Adrian R.

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3163969
Copyright 2005 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

© Adrian R. Willoughby

All rights reserved
2005

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

For my parents.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Contents

Dedication ii

List of Tables v

List of Figures vi

Chapter

I. Introduction 1

A Brief History of Cognitive Electrophysiology
The Error-Related Negativity
Four Theories of the ERN
Error Detection (The Mismatch Hypothesis)
The Conflict-Monitoring Theory
The Reinforcement Learning Theory
The Emotional Processing Theory
The ERN and Psychological Disorders and Lesions
Obsessive-Compulsive Disorder and Tourette's Syndrome
Schizophrenia
Parkinson's Disease
Alzheimer's Disease
Aging
Lesions
Research into the Anterior Cingulate Cortex
The ACC
The ACC and Action Monitoring
The ACC and Attention
The ACC and Emotion and Pain
The ACC in Animals
Testing Theories of ERN function

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

II. Experiment 1 51
Experiment 1a
Design and Rationale
Methods
Participants
Procedure
Electrophysiological Methods
Results
Behavioral Results
Electrophysiological Results
Experiment 1b
Design and Rationale
Methods
Participants
Procedure
Electrophysiological Methods
Results
Behavioral Results
Electrophysiological Results
Experiment 1 Discussion

III. Experiment 2 84
Design and Rationale
Methods
Participants
Procedure
Electrophysiological Methods
Results
Behavioral Results
Electrophysiological Results
Experiment 2 Discussion

IV. General Discussion 114
The Basic Results
Implications for Theories of the ERN
The Affective Theory
Conclusion

Bibliography 124

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Tables

Table

2.1 Association strengths for each outcome value in Experiment 1a and example Attneave shapes. 53

2.2 Association strengths for each outcome value in Experiment 1b and example Attneave shapes. 68

3.1 The association strengths in Experiment 2. 92

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Figures

Figure		
1.1	An error detection model of the ERN.	8
1.2	An illustration of the conflict-monitoring model.	14
1.3a	An illustration of hypothetical psychological processes in a response selection task.	17
1.3b	Illustration of possible psychological processes where two response representations are active.	17
1.3c	A different representation of the conflict-monitoring model.	18
1.4	Illustration of putative psychological processes and the measurements used to assess them.	20
1.5	The RL-ERN model.	22
2.1	A schematic representation of Experiment 1.	54
2.2	Experiment 1a change in accuracy over the block.	56
2.3	Experiment 1a accuracy for the different associations.	57
2.4	Experiment 1a accuracy for different confidence ratings.	58
2.5	Experiment 1a ERN averaged over all conditions.	60
2.6	Experiment 1a change in ERN over the block.	60
2.7	Experiment 1a ERN for the four confidence ratings.	61
2.8	Experiment 1a ERN for grouped confidence ratings.	62
2.9	Experiment 1a change in ERN over the block for the low and high confidence ratings.	63
2.10	Experiment 1a ignorance and violation errors.	65
2.11	Experiment 1a waveforms following presentation of value stimuli.	66
2.12	Experiment 1b change in accuracy over the block.	69
2.13	Experiment 1b accuracy for the different associations.	70

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.14 Experiment 1b accuracy for the different confidence ratings. 71

2.15 Experiment 1b ERN averaged over all conditions. 72

2.16 Experiment 1b change in ERN over the block. 73

2.17 Experiment 1b ERN for the four confidence ratings. 73

2.18 Experiment 1b ERN for grouped low and high confidence ratings. 75

2.19 Experiment 1b change in ERN over the block for the grouped low and high confidence ratings. 76

2.20 Experiment 1b ignorance and violation ERNs. 77

2.21 Experiment 1b waveforms following presentation of value stimuli. 78

3.1 A schematic illustration of Experiment 2. 91

3.2 Experiment 2 learning curves for the three association strengths. 93

3.3 Experiment 2 learning over the course of the block for the three association strengths. 94

3.4 Experiment 2 accuracy for the different confidence ratings. 95

3.5 Experiment 2 confidence for responses in the different association strengths. 96

3.6 Experiment 2 ERN averaged over all conditions. 97

3.7 Experiment 2 change in ERN over the block. 98

3.8 Experiment 2 change in ERN over the block for the strong association condition. 99

3.9 Experiment 2 change in ERN over the block for the weak association strength. 100

3.10 Experiment 2 change in ERN over the block for the random association strength. 101

3.11 Experiment 2 ERN for outcomes of different expectancy. 102

3.12 Experiment 2 change in ERN over the block for the different confidence ratings. 103

3.13 Experiment 2 ERN for stimuli of different frequency. 105

3.14 Experiment 2 ignorance and violation ERNs. 106

3.15 Experiment 2 waveforms following value presentation. 107

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.16 Experiment 2 waveforms for expected, better than expected and worse than expected values. 108

3.17 Error and correct waveforms from four individual subjects in Experiment 1b. 109

3.18 Error and correct waveforms from two individual subjects in Experiment 2. 110

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter I

Introduction

A Brief History of Cognitive Electrophysiology

Cognitive electrophysiology can boast a long and distinguished history that includes contributions from some of Europe's leading scholars. For many years, philosophers had contemplated the nature of thought and the mechanisms by which it operated. But it was Galen, a 2nd century Greek physician and philosopher (whose medical authority extended into the 16th century) who was perhaps the first to investigate experimentally the operation of the nervous system (in fact, the laryngeal nerve, a focus of his studies, is also known as Galen's nerve in his honor). Galen proposed that pneuma (breathe or spirit) was fundamental to all forms of life. This pneuma could take three forms, one of which was the pneuma physicon (the animal spirit – from the Latin, anima, or soul) that distinguished animals from other living organisms. Galen thought that this pneuma physicon existed in the nerves and was responsible for movement, sensation and thought. It was not until the late 18th century that evidence began accumulating that challenged this Galenist position.

Around 1780, Galvani began research into what he called 'animal electricity,' which culminated in the publication of 'De Viribus Electricitatis in Motu Musculari:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Commentarius' in 1791. He reported a number of experiments that showed electrical stimulation of the sciatic nerve in a frog resulted in muscle contraction and leg movement. Galvani concluded that animals possessed intrinsic electricity that was used to control motor behavior. However, he was unable to demonstrate electrical activity in the actual nerve itself. This crucial discovery had to wait until the middle of the 19th century, when Du Bois-Reymond, renowned for his skill in instrumentation and experimentation, demonstrated the propagation of electrical impulses down the nerve (Brazier, 1984).

Aware of Du Bois-Reymond's research, the English physician Richard Caton began his search for electrical activity in the brain. In 1875, he published what are regarded as the first reports of what are now known as Event-Related Potentials (ERPs) and the Electroencephalogram (EEG). Using a reflecting galvanometer as a means of amplification, he observed variations in the electrical currents of the brain when the retina was stimulated by light. Furthermore, he was able to localize the region of the brain that responded to visual stimulation to the contralateral visual cortex. While these ERPs were the major focus of Caton's research and commentary, he also noticed weak spontaneous electrical activity in the animal's brain even when no stimuli were being presented. It is this ongoing activity that became known as the EEG and has since received so much interest. Because Caton published his reports in medical journals that were not widely read in continental Europe, many other European scholars were unaware of them; therefore, shortly after Caton's discovery, Adolf Beck (in Poland), Ernst Fleischl von Marxow (in Austria) and Vasili Danilevsky (in Russia) all reported similar findings

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Brazier, 1988). It is perhaps because of these contemporaneous reports of brain electrical activity, and their intellectual legacy, that such a strong tradition in EEG research is held so widely throughout Europe.

The next major advance in cognitive electrophysiology came in Germany in the 1920s with Hans Berger's research into scalp recordings of electrical brain activity in humans. This culminated in 1929 with the first published human EEG recordings, from Berger's son, Klaus. Berger's own research focused on rhythmic activity in the EEG (that he termed alpha waves, which are still studied today). He also coined the term 'electroencephalography' and is considered, perhaps quite aptly, the father of human electrophysiology.

The 1950s and 1960s saw two different revolutions, both of which affected cognitive electrophysiology. The computer revolution allowed the wealth of electrical activity produced by the brain (the EEG) to be recorded and analyzed much more efficiently (e.g., Carrie, 1972; Estrin, 1965; Giannitrapani et al., 1971; Ponsen, 1978). Furthermore, amplification and recording techniques and hardware advanced to enable higher quality recordings. However, the essentials of EEG recording remained as they were when Berger first reported them. The cognitive revolution occurred with (and was, in part, motivated by) the computer revolution. Scientists became increasing interested in mental processes and those parts of the brain responsible for them. It was quickly realized that ERPs or changes in rhythmic patterns in the EEG might index psychological processes. And so the investigation into which psychological phenomena accompanied changes in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

electrical brain activity and which changes in brain electrical activity accompanied psychological phenomena began in earnest. Cognitive electrophysiology and newer neuroimaging techniques, such as PET and fMRI (developed to localize brain regions involved in performing mental functions) are still central in the field of cognitive neuroscience, which is at the forefront of modern research into the mind and brain.

The Error-Related Negativity

Rightly, much of psychology is devoted to the investigation of the complex chain of processes involved in the selection and performance of actions. This includes many cognitive operations, from being motivated to initiate an action, to identifying stimuli, attending to the appropriate one, retrieving relevant information from memory and selecting the correct response. However, for survival in their environment, it is essential for animals to monitor their on-going actions. This monitoring enables correction of any mistakes, adjustments to the animal's actions to prevent further mistakes from being made or a change in strategy if the desired results are not being achieved. Therefore, it is equally important to study how cognitive processes are monitored and adjusted to ensure successful interaction with the environment, if we are to have a fuller understanding of mental operation.

Some of these monitoring processes – how errors are detected and how they might be corrected – became amenable to investigation following the serendipitous discovery of what became known as the Error Negativity (Ne; Falkenstein et al., 1991) or Error-Related Negativity (ERN; Gehring et al., 1993). This is a particular pattern of electrical

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

brain activity recorded at the scalp that appears to be strongly related to the commission of errors and is characterized by a negative deflection in the response-locked ERP peaking at medial-frontal electrode sites approximately 100 ms after an erroneous response, lasting for about 150 ms.

Subsequent research has uncovered a number of consistent experimental effects on the ERN. These include results showing it occurs following many different types of errors and for different stimulus and response modalities (e.g., Falkenstein et al., 1991; Holroyd et al., 1998; Miltner et al., 1997; Nieuwenhuis et al., 2001; Van't Ent and Apkarian, 1999), it is smaller following slow errors and larger following small or corrected errors (e.g., Gehring et al., 1995) and it is smaller in older adults (e.g., Falkenstein et al., 2001; Gehring et al., 1995; Nieuwenhuis et al., 2002). Furthermore, preliminary results suggest that it can be seen in participants who observe another person making an error (van Schie et al., 2004). Electrical source analysis suggests the ERN is generated by a neural source in the medial-frontal brain region, usually located in the anterior cingulate cortex (ACC; e.g., Dehaene et al., 1994; Miltner et al., 1997). Also, repetitive transcranial magnetic stimulation (rTMS) appears to reduce the amplitude of the ERN when it is applied over medial-frontal regions compared with lateral ones (Rollnick et al., 2004). Intracranial recordings from depth electrodes (Brazdil et al., 2002) also suggest the ERN may be generated by a source in the ACC; however, it also appears that a wider system may be involved, including medial-temporal, orbito-frontal and dorso-lateral prefrontal areas. Further research inspired by the ERN has uncovered several ERN-like components that may be functionally related to the ERN itself. A similar negativity, also maximal at

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

medial-frontal electrode sites, has been observed following error feedback (e.g., Holroyd and Coles, 2002; Miltner et al., 1997). This has become known as the "feedback ERN" to distinguish it from the original "response ERN" but is frequently simply referred to as the ERN. While these two components share the same name, occur under similar circumstances and have similar characteristics, it is still an open question whether they reflect the same underlying psychological process.

A further ERN-like component has been observed following monetary losses in a gambling task that appears independent of response accuracy (Gehring and Willoughby, 2002). This component occurs following negative feedback and shares many similarities with the feedback ERN; however, it was given the more generic name, the "Medial-Frontal Negativity" (MFN), because it does not seem to be, strictly speaking, error-related. The term MFN has also been used as a wider categorical term to which all the ERN-like phenomena may belong. Again, while sharing similarities with each other, it is an open question to what extent the ERN and MFN reflect the same processes. In fact, one report (Gehring and Willoughby, 2004) claims that while similar, the MFN and the response ERN are characterized by different scalp distributions, suggesting the underlying sources have, at least partly, a different spatial configuration. As yet, no experiment has been reported that directly compares the feedback ERN and the MFN.

Smaller ERN-like activity has also been observed following correct responses (sometimes called the "Correct-Related Negativity," or CRN; e.g., Vidal et al., 2000). Again, there is debate concerning whether this activity is functionally equivalent to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ERN and whether it is generated by the same regions (e.g., Falkenstein, 2004; Luu and Tucker, 2001); however, it is potentially an indication that the same processes are executed on both error and correct trials.

The earliest interpretation of the ERN was that it reflects an error detection or correction process. While this is still a predominant view, several other theories have also been proposed. These include the conflict-monitoring theory (e.g., Botvinick et al., 2001), the reinforcement learning theory of the ERN (RL-ERN; Holroyd and Coles, 2002) and a more general affective evaluation theory (e.g., Luu et al., 2003).

Four Theories of the ERN

Error Detection (The Mismatch Hypothesis)

The history of the error detection theory can be traced back to the beginning of ERN research (e.g., Falkenstein et al., 1991). There are two possible ways in which the ERN could be associated with error detection, it could be a manifestation of the error detection process itself or it could reflect the outcome of the error detection process. These two possibilities generate different predictions about the amplitude of the ERN for different types of error. For example, an error that is easy to detect (one that is very different from the correct response) might engage the error detection system only minimally and result in a small ERN. However, the output of the error detection system for such an error might indicate a significant mismatch between the correct and error representations, so the ERN might be large. Figure 1.1 shows a simple illustration of these two possibilities; either the ERN reflects the output of the process, or the process itself.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

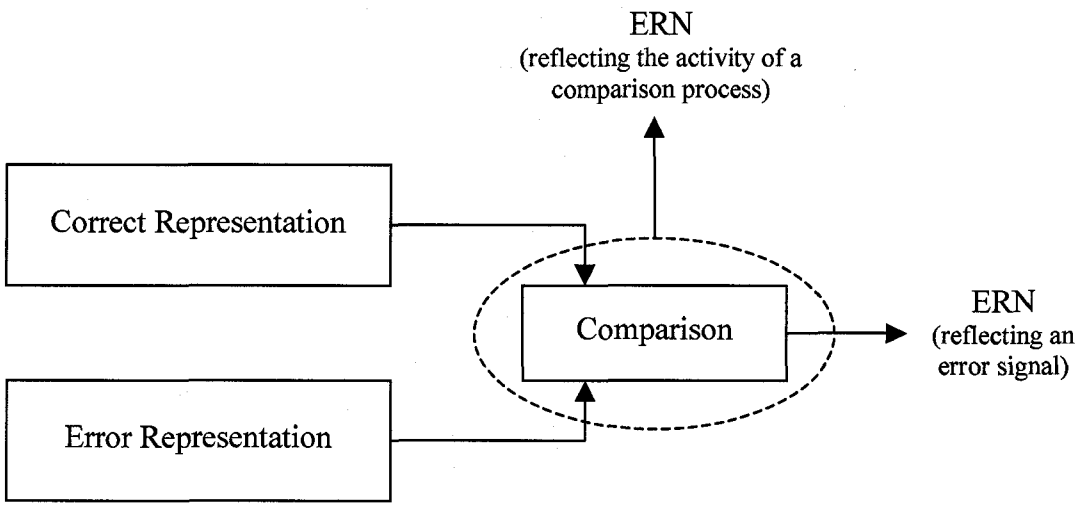


Figure 1.1 An error detection model of the ERN. The ERN could reflect the comparison process itself or the output of such a process.

Falkenstein et al. (1991) suggested the ERN reflected the comparison process between the erroneous response and the (correct) outcome of the completed response selection process. Coles et al. (2001), on the other hand, proposed that the ERN was the outcome of a comparison between the representation of the error response (derived from an efference copy of the motor output command) and the representation of the correct response (derived from on-going stimulus processing). Evidence suggests proprioceptive feedback from the erroneous movement cannot be used in the comparison because the ERN occurs too soon for feedback from the limb to be generated and transmitted back to the comparison system (its onset is approximately 100 ms following the beginning of the response, measured using EMG; see Gehring et al., 1993).

Several lines of evidence are consistent with the view that the ERN reflects an error detection process. Berstein et al. (1995), employing a four-choice reaction time task, found the ERN varied with the size of the difference between the correct and incorrect

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

responses. However, using a similar procedure with hands and feet as response effectors rather than fingers (because of the neural and muscular interaction between fingers), Gehring and Fencsik (2001) found the opposite result – that the ERN was largest for similar errors (rather than different ones). This suggests the ERN might reflect activity of the comparison system rather than its output because the ERN is larger during a task in which the comparison system is more engaged.

Further evidence consistent with the ERN reflecting an error detection process comes from experiments that manipulate the correct response representation, on which the comparison relies. Scheffers et al. (1999) used a number of different manipulations to compromise the correct response representation, including increasing load in visual and memory search tasks, depriving participants of sleep and increasing variability in stimulus-response (S-R) mapping. The amplitude of the ERN was found to be decreased whenever the correct response representation was compromised. A similar demonstration of the importance of the correct response representation comes from Scheffers and Coles (2000). An Eriksen flanker task with degraded stimuli was used to influence participants' subjective beliefs about their performance. The results showed that the ERN was larger when the participants thought they had made a mistake (because of an incorrect response representation) than when they thought they were correct, regardless of their actual accuracy.

Similarly, the incorrect response representation can be manipulated, although only one technique has been used. Rather than allow participants immediate access to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

accuracy of their performance, some experiments have delayed error feedback – therefore, delaying generation of the incorrect response representation. For example, Miltner et al. (1997) had subjects perform a time estimation task. Because participants were unable to determine their own accuracy, they had to wait for feedback given through a variety of stimulus modalities (auditory, visual and somatosensory). The results demonstrated that error feedback evoked an ERN-like negativity (the feedback ERN). Furthermore, not only was the putative ACC generator of the negativity almost identical for all modalities, but also very similar to those reported for the response ERN. More recent studies (e.g., Badgaiyan and Posner, 1997; Butterfield and Mangels, 2003; Ruchsow et al., 2002) have resulted in similar findings.

These results suggest that the ERN reflects a comparison of representations regardless of whether the representations are themselves correct and if one representation is not immediately available, the comparison process must wait until it becomes so. While these experiments show both correct and incorrect response representations are important for the ERN to be elicited, they are merely consistent with the error detection theory rather than strong evidence for it. The ERN might reflect downstream activity that relies on both representations and the comparison between them, such as the implications of making an error.

In fact, results from studies that have closely examined the time course of errors, corrective actions and the ERN suggest this may indeed be the case. Gehring and Fencsik (1999) showed that the ERN began after corrective actions had already been

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

initiated. Similarly, Rodriguez-Fornells et al. (2002) found that components of the Lateralized Readiness Potential (LRP) that reflected error correction preceded the ERN. These findings suggest that the ERN occurs too late to reflect error detection directly. Another potential problem for the error detection theory of the ERN is the discovery of the MFN following monetary losses. The original finding (Gehring and Willoughby, 2002; since replicated by Nieuwenhuis et al., 2004) showed that the amplitude of the negativity following feedback in a gambling task was sensitive to gaining or losing money rather than the accuracy of the response. These results suggest that the negativity could be related to processing utility rather than accuracy information and, to the extent to which the ERN and MFN reflect the same psychological process, that the ERN could be explained in terms of processing the negative consequences of making a mistake.

The observation of the CRN also argues against the ERN being simply a reflection of an error detection process (if, indeed, it is the case that the ERN is an enhanced CRN rather than a separate component). While the CRN has been noted in a number of experiments (e.g., Falkenstein et al., 2000), it has rarely been the target of explicit research. However, Vidal and his colleagues are the notable exception to this. Using a number of choice reaction time tasks they demonstrated both a traditional ERN on error trials and a CRN on correct trials (Vidal et al., 2000). Therefore, they proposed the ERN cannot represent merely an error detection process but suggested, instead, that it reflects a response evaluation or an emotional process.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In defense of the error detection theory, Coles et al. (2001) suggested several reasons why ERN-like activity might be seen on correct trials, including participants thinking they had made an error or using different subjective accuracy criteria (e.g., response time), partial errors that are corrected before the actual response is made and on-going stimulus related processing. Vidal et al. (2003) rejected the claim that the CRN could be due to participants mistakenly thinking they had made an error. One of the experiments on which the CRN was observed was a very simple Go/No go task with unambiguous stimuli in which response accuracy was immediately clear. Furthermore, with a number of re-analyses of their original data (such as plotting and grouping responses as a function of reaction time) they found no evidence that participants were using subjective accuracy criteria different from those imposed by the experimenter. The stimulus-locked waveforms following correct responses were also examined and no evidence was found of a stimulus-locked component contaminating the response-locked waveforms that could account for the CRN. Luu and Tucker (2001) also found the scalp distribution and source localization for the negativity on both correct and error trials to be very similar, suggesting the psychological mechanism underlying both of them is the same, which is unlikely to be the case if the CRN is an artifact of stimulus related processing. However, the CRN is typically seen most clearly using current source density analysis techniques that are particularly sensitive to superficial sources, which suggests their neural generators may, in fact, be different (Falkenstein, 2004).

One consequence of the observation of the CRN was a modification of the error detection theory to propose both a comparison process and an error detection process (e.g.,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Falkenstein et al., 2001). This dual process theory is designed to account for both the CRN (the comparison process that occurs on both error and correct trials) and the ERN (the error detection process); however, opinion is divided as to whether they reflect the same psychological processes and share similar scalp distributions and neural sources.

The Conflict-Monitoring Theory

The conflict-monitoring theory was inspired by functional magnetic resonance imaging (fMRI) research showing increased ACC activity under conditions of high response conflict (e.g., Carter et al., 1998; Carter et al., 2000) and was developed to help explain some of the results that were problematic for the error detection theory (e.g., Botvinick et al., 2001). The logic behind this theory is that errors typically occur when two or more mutually incompatible response representations are active; and it is the conflict between them (their mutual inhibition) that is detected rather than the error itself. Following an incorrect response (caused by activation of the incorrect response representation), continued stimulus processing leads to activation of the correct response representation; therefore, the conflict between the simultaneously active error and correct representations peaks shortly after the error is made, when the ERN is observed. Figure 1.2 shows an illustration of the model derived from the conflict-monitoring theory.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

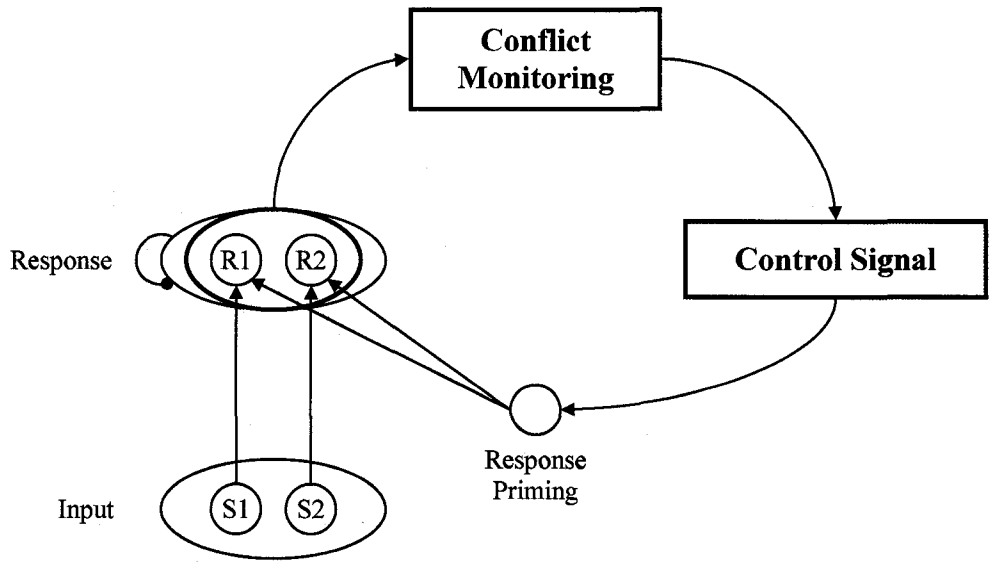


Figure 1.2 An illustration of the conflict-monitoring model. From Botvinick et al. (2001).

One of the benefits of the conflict-monitoring theory is that it gives a simple explanation of why a CRN might be seen on correct trials; two competing responses might generate response conflict but the correct response might still be selected. Carter et al. (1998), using a continuous performance task, showed increased ACC activity not only on error trials but also on correct trials associated with higher response conflict. In fact, increased ACC activity under conditions of high response conflict is a robust finding in fMRI research over a number of experimental paradigms (e.g., Barch et al., 2000; Nelson et al., 2003).

However, on closer inspection, the time course for conflict monitoring on correct and error trials should be different. Conflict is resolved before the response for correct trials but peaks after the response for incorrect trials (e.g., Yeung et al., 2004). While this kind of prediction is hard to investigate using fMRI because of its low temporal resolution, it is amenable to electrophysiological investigation. A candidate ERP for conflict

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

monitoring prior to a correct response is the N2. This ERP has been shown to be enhanced on no go trials of the Go/No go task. The traditional interpretation of the enhanced N2 is that it reflects inhibition of a prepotent response (e.g., Jodo and Kayama, 1992). However, it has been argued that the no go N2 could actually reflect the monitoring of conflict between the go and no go response representations in an experiment that investigated the N2 and ERN in a traditional Flanker task (van Veen and Carter, 2002). The results showed that the N2 following stimulus presentation on correct trials increased with response conflict. Furthermore, the N2 and ERN had similar scalp distributions and the results of source dipole modeling generated almost identical ACC sources. However, it is not unanimously accepted that the no go N2 and ERN are manifestations of the same underlying cognitive function. Falkenstein et al. (1999) showed that the no go N2 and ERN responded differently to different experimental manipulations. The no go N2 was larger both in better performing participants and for visual than auditory stimuli, while neither of these manipulations was found to effect the ERN. Furthermore, they showed that the N2 had a more frontal scalp distribution than the ERN. However, Nieuwenhuis et al. (2004) suggested a methodological artifact involving the relative auditory and visual similarity of the stimuli may have caused these effects. They went on to show an enhanced no go N2 in both the visual and auditory modalities whenever the go and no go stimuli were similar and consequently suggested the no go N2 and ERN did, in fact, reflect the same process.

While the conflict-monitoring theory might be able to explain ACC activation on correct trials in fMRI experiments and an enhanced N2 in electrophysiological studies, it still

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

fails to explain the observation of the CRN. If conflict is resolved on correct trials before the response, further conflict related activity following the response should not be observed.

Further concerns surrounding the conflict-monitoring theory involve the assumption of conflict as a psychological process and its measurement. Response conflict is defined as "the simultaneous activation of mutually inhibiting units" (Botvinick et al., 2001, p. 630) and it is calculated operationally by the model as the Hopfield energy in the response layer (that is, the total activity of the units in the response layer of the network weighted by the inhibitory connections between them). This definition captures the intuitive assumptions of response conflict; that there are two (or more) simultaneously active representations and that they are mutually inhibitory. While these are reasonable assumptions of a neural architecture, they need not necessarily be true. It might be the case that there can only be one response representation active at any one time, or that there could be any number active but they are not mutually inhibitory (indeed, the reinforcement learning model of the ERN discussed later does not include inhibition between response representations). Figures 1.3a and 1.3b schematically illustrate two different types of response system involving one or two response representations that do not necessarily involve response conflict; Figure 1.3c shows the conflict-monitoring model in a similar format. The comparison between correct and incorrect response representations in the error detection theory proposed by Coles et al. (2001) is shown.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

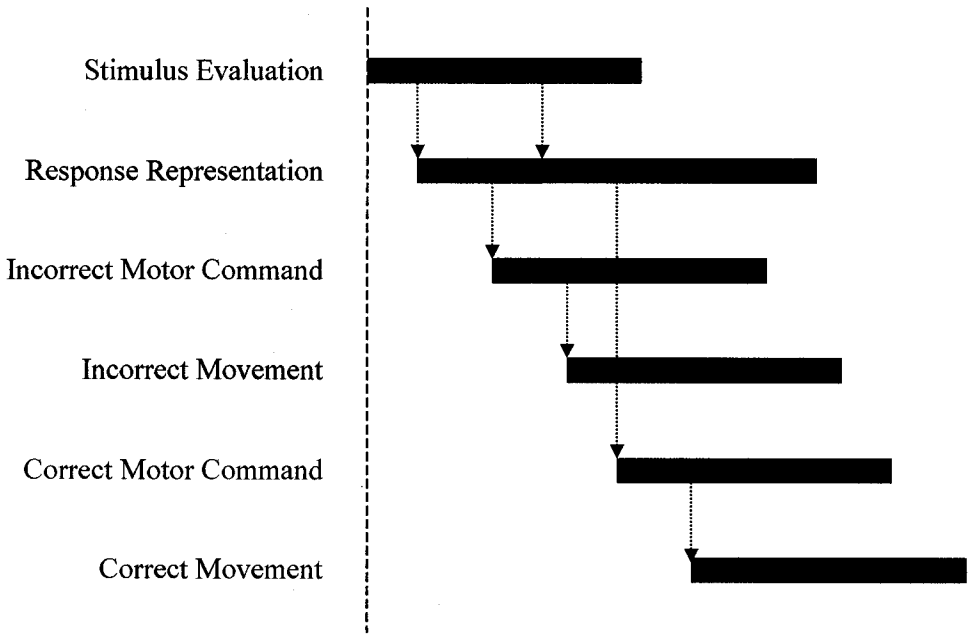


Figure 1.3a An illustration of hypothetical psychological processes in a response selection task. Horizontal bars indicate ongoing putative psychological processes. Vertical arrows indicate output to the next process. Only one response representation is ever active. Green indicates correct processes, red incorrect ones.

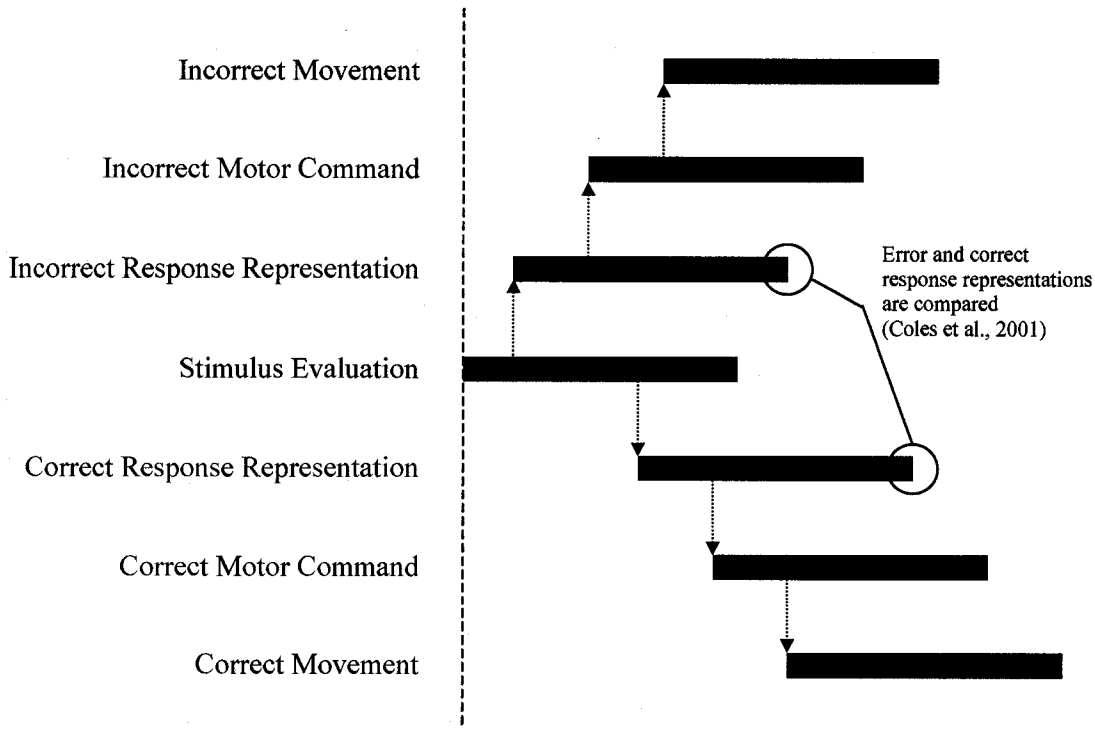


Figure 1.3b Illustration of possible psychological processes where two response representations are active.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

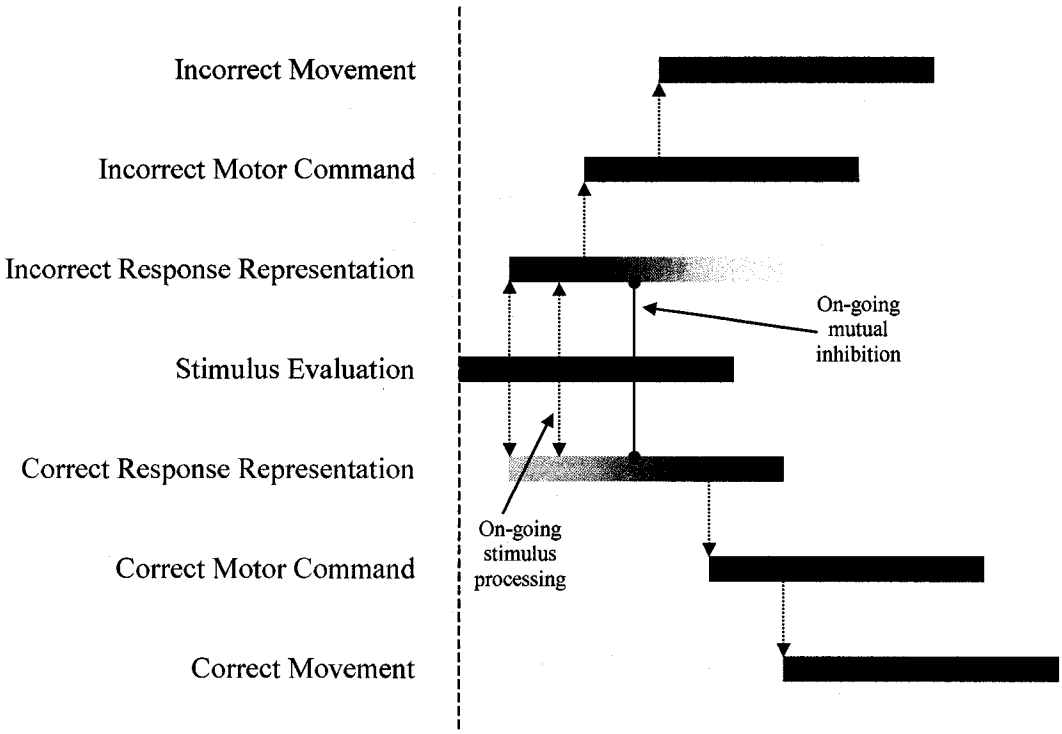


Figure 1.3c A different representation of the conflict-monitoring model. Mutual inhibition decreases unsupported activity quickly.

The notion of response conflict has been used in psychological explanations and models for many years (e.g., Berlyne, 1957) but deriving an actual measure of response conflict might not be as simple as it first appears. The traditional behavioral measures, reaction time and error rate, have been suggested to index response conflict such that increased conflict will lead to an increase in either reaction time or error rate (or both). However, manipulations that are designed to increase response conflict frequently alter other processes as well as the hypothesized response selection process; therefore, it is hard to tell whether the increased reaction time or error rate actually resulted from prolongation of the response selection process itself (and concomitant increase in the degree of response conflict) or by increasing the duration of some other process. Furthermore, if a task activates more than one response, top-down control is likely to be recruited to inhibit

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the inappropriate responses, which might also affect other cognitive processes. For example, in the congruent condition of the flanker task stimulus identification is simple because the target and flanker stimuli are the same and are associated with the same response. However, in the incongruent condition, the increase in stimulus complexity might have several consequences. These might include an increase in the duration of the stimulus identification process or increased recruitment of control processes to inhibit the incorrect response. None of these are necessarily related to response conflict. Another way of conceptualizing the difference in conditions involves considering the rules by which the participants are operating. While the experimental rules are identical in both the congruent and incongruent conditions – 'respond to the central stimulus and ignore the flanking stimuli' – the actual rules used by the participants might be quite different and be associated with different processes. For example, in the congruent condition the internalized rule might be, 'if all the stimuli are the same respond appropriately to any one of them as soon as possible,' while for the incongruent condition the rule might be, 'if the central stimulus is different focus attention on the central stimulus, identify it and respond appropriately on the basis of that stimulus alone.' The rule used for the incongruent condition requires engagement of an attentional process that the rule for the congruent condition does not.

Measures other than traditional reaction time and error rate have also been used to try to assess response conflict. They are all based on the same assumption and essentially measure the same process at different points, that is, erroneous activity on correct trials (also known as partial errors). The assumption is that on high conflict trials, an erroneous

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

response is likely to be prepared and may be carried out at a sub-response threshold level. Therefore, experiments have used the LRP to measure the degree of erroneous motor preparation, electromyography (EMG) to measure the degree of erroneous muscle activity and force output to measure the degree of erroneous response. A major question then, is whether mere co-activation of different response systems necessitates conflict between their representations during the (prior) response selection process. While this might make intuitive sense, it need not necessarily be the case. Figure 1.4 shows an illustration of the processing stages at which these various measures are taken.

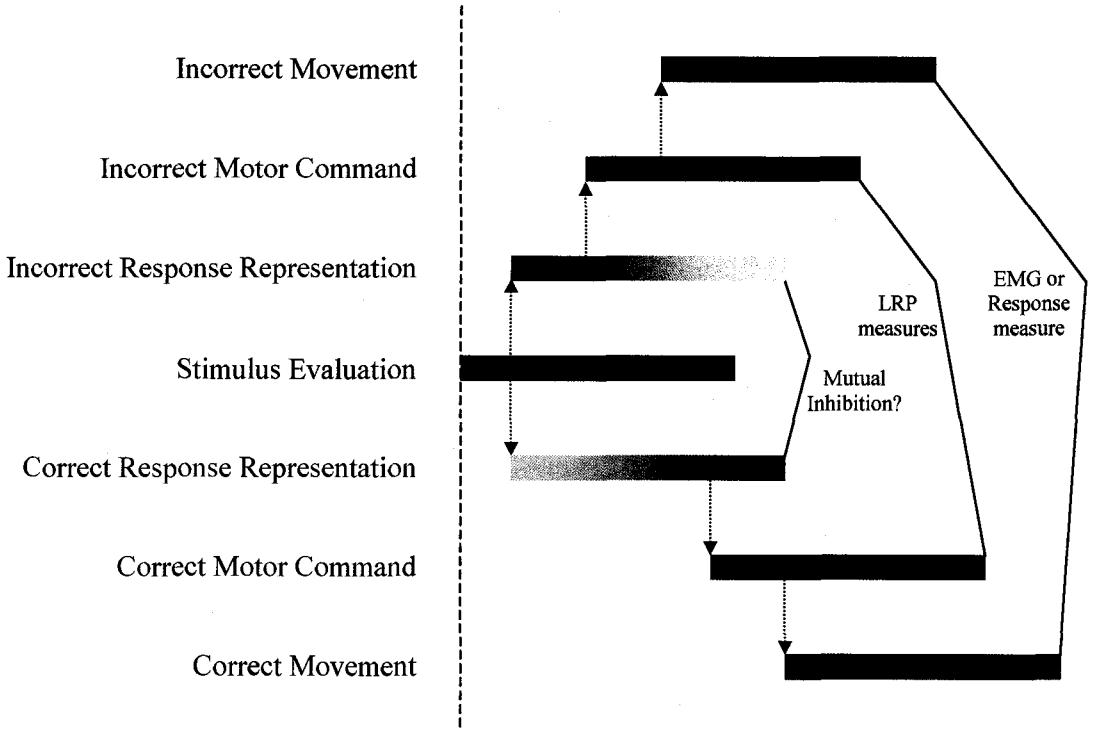


Figure 1.4 Illustration of putative psychological processes and the measurements used to assess them.

Furthermore, as noted by Luu et al. (2000) and Gehring and Fencsik (2001), using the LRP as a measure of conflict may be problematic. The LRP is a difference measure, therefore high levels of activity in the motor cortex of both hemispheres (resulting in a high degree of putative response conflict) could result in a low LRP value. In fact, in a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

study investigating this, Rodriguez-Fornells et al. (2002) found very similar LRP waveforms for an erroneous response followed by a corrective one and a correct response that consisted of the same physical movements (a button press with a finger on one hand followed by another with a finger on the other hand). This suggests that, when physical movement is controlled for, the LRP is the same in conditions where conflict is believed to be high (when an error is quickly followed by a corrective response) and where conflict is believed to be low (on a correct response).

The Reinforcement Learning Theory

The Reinforcement Learning model of the ERN (RL-ERN) proposed by Holroyd and Coles (2002; Figure 1.5) was developed from the original error detection model and, therefore, shares some similarities with it. However, instead of detecting errors, it is proposed that the system monitors for events that are worse than expected (of which errors are a subset). Then, rather than (or, perhaps, in addition to) inducing immediate compensatory changes or recruiting higher level strategic control mechanisms, the model proposes that the system learns both to improve task performance and to update its expectations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

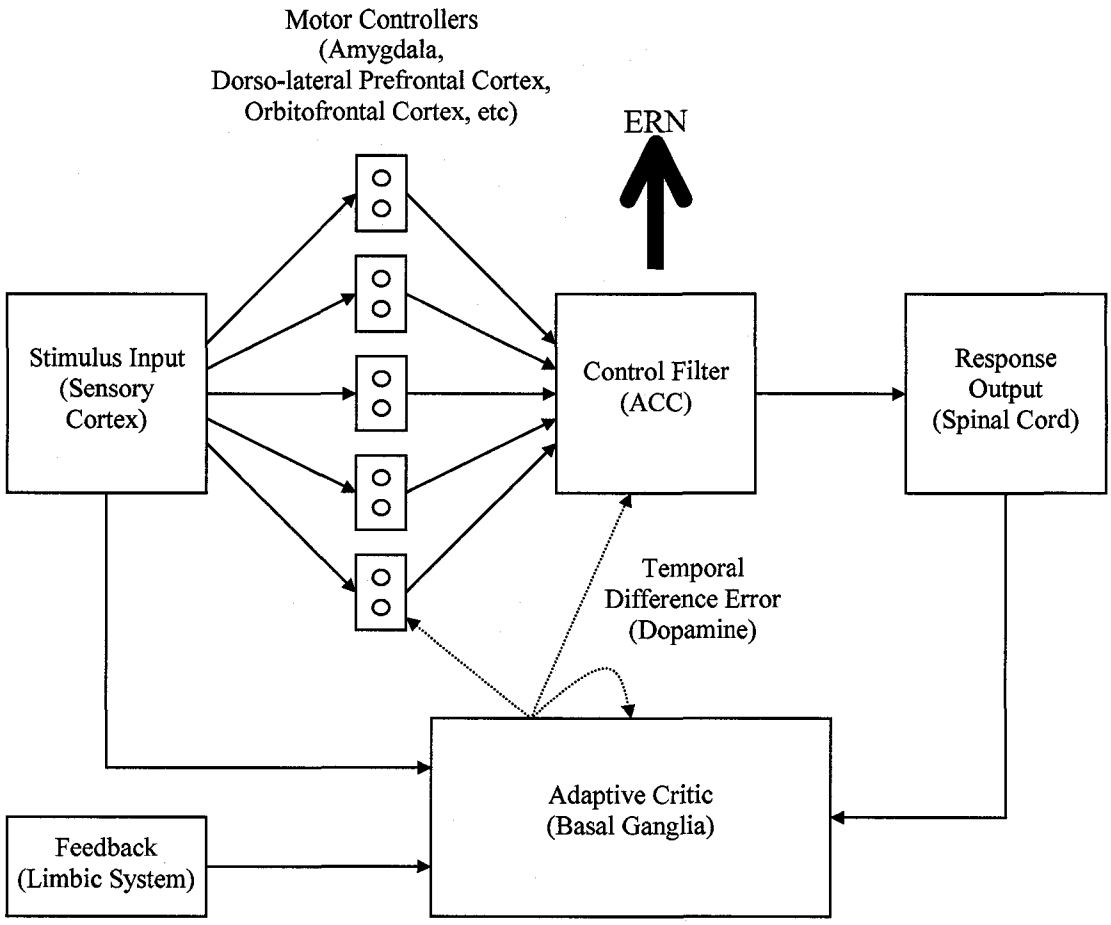


Figure 1.5 The RL-ERN model. Adapted from Coles and Holroyd (2002).

An adaptive critic (or comparator) receives information about on-going events and based on this input makes predictions about what the outcome of the events will be. These predictions are then compared with the actual outcome. If the prediction is wrong, a temporal difference error signal is generated. This error signal enables the adaptive critic itself to update its predictions, it allows a control filter (that selects which motor controller will control the motor system and, therefore, actual output) to learn to which motor controller to give authority and it teaches the motor controllers themselves which response to select. It is claimed the ERN is caused by the effect of the temporal difference error at the control filter (the ACC) on its arrival.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The RL-ERN model shares some similarities with other theories of ERN function. Both the RL-ERN model and the error detection theory postulate the ERN results from a discrete comparison process (that takes place when two representations are available), whereas the conflict-monitoring theory proposes the ERN is a particular event during a continuous monitoring process. However, while the comparison in the error detection and conflict monitoring theories is between response representations (the response they represent is compared in the error detection theory and their relative activity is compared in the conflict-monitoring theory), the comparison in the RL-ERN is between predicted and actual outcomes (which includes, but is not limited to, responses).

The RL-ERN approach is a sophisticated and ambitious attempt to combine two different fields, the electrophysiology of action monitoring and reinforcement learning theory, into a unified model. Furthermore, because cognitive processes have been mapped onto neural structures, it integrates cognitive and biological psychology. Therefore, it is possible then to judge the model on separate grounds – its standing as a cognitive model of the putative psychological processes underlying the ERN, its neurophysiological plausibility and the proposed mapping of function onto structure. Also, given the novel conception of this model, it enables new predictions to be tested, such as how the ERN might change during the course of learning.

Considering the RL-ERN model solely on cognitive grounds, several issues arise concerning the redundancy of response representations and the response selection

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

process. The current instantiation of the model has five motor controllers and therefore five sets of response representations. While the nature of the response representations in the model is the same (all are units representing a manual button press), the actual conceptualization of them is somewhat different. The theory describes them as not simply representing pressing one button or another (as in the model) but more complex motivations to act in certain ways; for example, to approach or avoid the stimulus, or to select a stimulus that will lead to short-term gains regardless of the long-term effects. The relationship between these types of representations and the motor commands to execute the selected actions may be complicated and vary between the different motivation systems and is not considered in this model. The response selection mechanism, while described in more detail by the RL-ERN model than many others, is of a slightly unusual nature. It is a two-part process whereby each motor controller selects one of the potential responses and the control filter selects one of the motor controllers. Most psychological models would simply assume that one response representation would be activated, which would then initiate an actual response.

In order to avoid the response representation redundancy and two-stage response selection process (and, perhaps, consider the model in more familiar terms), a model in which a number of systems influencing central response representations could be hypothesized; the control filter could then exert its effect by modulating the influence of the systems on a single set of response representations. Furthermore, it could be possible for the systems to interact and this interaction could actually be the instantiation of the control filter. For example, if two incompatible response representations (e.g., approach

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and avoidance) were being activated, connections between the systems could reduce the influence of the inappropriate systems so only the most appropriate one would activate the response representation, without the need for an actual control filter. This kind of interaction is not modeled explicitly in the RL-ERN but is alluded to in the description of the motor controllers as 'acting semi-independently' (Holroyd and Coles, 2002, p. 685).

Reinforcement learning models depend on the temporal difference signal for learning to take place, which Holroyd and Coles propose is transmitted by the mesencephalic dopaminergic system. Therefore, a prediction of the RL-ERN is that if these dopaminergic pathways were damaged, reinforcement learning could no longer be carried out. However, Berridge and Robinson (1998) have shown that rats with depleted striatal dopamine are still capable of reinforcement learning. Another issue (also discussed by Berridge and Robinson) concerns measurement of expectation (or prediction). A prediction can vary from a simple classically conditioned association to a more complex explicit cognitive expectation. This cognitive expectation involves more factors than just the probability of co-occurrence of the stimulus and the outcome, such as affect, incentive motivation and the semantic information in the representation of the predicted outcome. An example of this might be the 'gambler's fallacy.' Following a string of reds most people would bet black because they expect it is 'due' to come up, given their assumptions about the nature of strings of random events. However, given a truly random input to a pure learning system (like the RL-ERN), this expectancy would not be present because it would extract the true, underlying probability structure. In experiments on humans, explicit expectation is usually tested, which may have more

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

contributing factors than the expectation generated by the internal reinforcement learning system. This, then, leads to problems in knowing the actual nature of the expectation being measured, how to measure the expectation in which you are actually interested, how the measured expectation is related to the expectation generated by the reinforcement learning model and what the effect of violating the expectation is likely to be.

The RL-ERN model has received mixed empirical support from the few studies in which its predictions have be tested. In their original test of the model, Holroyd and Coles (2002) reported an experiment that examined the activity of response and feedback ERNs over the course of learning. The results showed that when the S-R mapping was random the feedback ERN was largest, but when the mapping was consistent the response ERN increased as it was learnt. This suggests the first event to give information about the accuracy of the response elicits the ERN. A similar experiment (Nieuwenhuis et al., 2002) included more S-R mapping conditions and tested both old and young participants. A crucial finding was that the feedback ERN was largest when subjects received error feedback they weren't expecting (i.e., information that on-going events were worse than expected). This was, however, only true for the young subjects.

In contrast to this supportive evidence, there exist a number of results that suggest the feedback ERN is not larger following unexpected error feedback. A study by Horst et al. (1980) that was designed to investigate the effects of subjective probability on the P300 component also reveals effects on the ERN. The primary results showed that the P300

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

varied with subjective probability; that is, subjectively unexpected events elicited a larger P300 than expected ones. However, while the study was not designed to examine the ERN explicitly, inspection of the waveforms suggests a feedback ERN was present following error feedback and, furthermore, it appeared to be the same both when the participants expected error feedback and when they thought they were correct (although it overlapped a larger P300 when the error feedback was unexpected).

In a study investigating ERPs for compound cues in a probabilistic learning task, Kopp and Wolff (2000) showed that in a high cue validity condition (where participant expectancy is high) the P3a component was sensitive to violations of expectancy. In the low cue validity condition (where expectancy is low), however, error feedback was followed by a larger N2, which displayed remarkably similar characteristics to the feedback ERN (in fact, Kopp and Wolff themselves noted the component could have been the ERN).

Butterfield and Mangels (2003) investigated violations of expectancy in a different way. Instead of using different association strengths between stimuli and responses, as in Kopp and Wolff (2000), they measured participants' confidence in their answers to trivia questions; it was assumed participants would least expect error feedback following answers given with high confidence. While they found in one version of the study (and not another) an effect of confidence on ERN amplitude, this was only after correcting for potentially overlapping later components. However, a much more pronounced result was an increase in P3a amplitude following error feedback after responses given with high

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

confidence (and, therefore, decreased expectancy of error feedback). This mirrors Kopp and Wolff's findings – that increased violation of expectancy results in an increase in the P3a component, not the feedback ERN.

The Emotional Processing Theory

Given results observed by Gehring and Fencsik (1999) and Rodriguez-Fornells et al. (2002) suggesting the ERN occurs too late to reflect an error detection process, speculation has intensified on what subsequent processing could be generating the ERN. One conjecture that has been present since some of the earliest ERN research (e.g., Gehring et al., 1993), but never modeled, is that the ERN might reflect the affective consequence of making a mistake. While a number of researchers have suggested this explanation (e.g., Gehring and Willoughby, 2002; Vidal et al., 2000), Tucker and his colleagues (e.g., Luu and Posner, 2003; Luu and Tucker, 2001) have been its major proponents. They suggest that the ERN, rather than reflecting a discrete neural event, is a manifestation of a synchronization of on-going theta activity that represents coordination between a number of limbic networks in action monitoring. Evidence for the emotional processing theory comes from a number of studies that have shown effects of motivational manipulations on the ERN (e.g., Gehring et al., 1993; Pailing and Segalowitz, 2003).

While the exact nature of the putative emotional processing system involved in the generation of the ERN has not been clearly outlined, there has been some consideration of the affective system and its relationship to the cognitive system. Yeung (2004)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

suggests we must consider both the input to the affective system, which presumably comes from the cognitive system and the function of the affective system, which is presumably to motivate us to behave in evolutionarily adaptive ways. For example, if we see a fearful stimulus (which is performed by the cognitive system), activity in the affective system motivates us to avoid it (there is also a simultaneous subjective feeling of fear, although it is unclear whether the feeling itself is the motivating factor). The questions Yeung raises are whether and to what extent we can separate the cognitive and affective systems and whether we can demonstrate that the ERN is generated by one system and not the other. Yeung proposes two possibilities; the first is that the cognitive and affective systems are empirically separable but the affective system uses input from the cognitive system (the weak dependence hypothesis), the second is that they are not empirically separable (the strong dependence hypothesis). If we accept the weak dependence hypothesis, we are proposing the cognitive system detects an error and activates the affective system, which motivates us to change our behavior to commit fewer errors (and results in a subjective negative feeling). If we accept the strong dependence hypothesis, we are proposing that detecting a mistake in and of itself motivates us to make fewer errors (and, again, results in a subjective negative feeling).

Two recent papers (although not comparing cognitive and affect processing) serve as an illustrative example of the distinction between a processing system and the input to it. Ridderinkhof et al. (2002) showed that ERN amplitude was diminished following alcohol consumption and suggested this was indicative of a compromised action monitoring system. Holroyd and Yeung (2003), using the RL-ERN as the compromised system,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

confirmed that this was indeed a possibility. However, they also showed that the results could be accounted for by degraded input to the conflict monitoring system. Therefore, there are two possibilities that have to be considered: one is that the action monitoring system is compromised, the other is that the input to that system is compromised but the system itself is unaffected. Analogously, changes in the ERN might not result from manipulations influencing the affective system but ones that influence its input from the cognitive system.

Many different types of psychological studies have provided evidence for relatively separate neural regions subserving cognitive and affective processing. Neuroimaging studies have shown different brain regions become active for cognitive and affective processing and begun characterizing their interactions (e.g., Drevets and Raichle, 1998). For example, Schneider et al. (1997) showed increase amygdala activity during mood induction. Neuropsychological studies have shown that lesions to different brain areas selectively impair cognitive or affective processing. For example, lesions to the amygdala result in what is now known as Kluver-Bucy syndrome, which is characterized by a number of inappropriate behaviors that seem to result from ascribing the wrong motivational properties to objects, such as trying to eat inedible objects or inappropriate sexual behavior. Furthermore, amygdala lesions can lead to an abnormal lack of fearful behavior in the face of threatening situations (e.g., Aggleton, 1992; Aggleton, 1993). Intracranial stimulation studies have shown that stimulation to different brain regions results in cognitive or affective behavior. For example, stimulation of sensory cortex results in sensations coming from the corresponding area on the body and stimulation of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the motor cortex results in involuntary movements of the corresponding limb (e.g., Penfield and Boldrey, 1937). Stimulation of the amygdala, however, results in fear behavior such as freezing (e.g., Davis and Whalen, 2001; De Molina and Hunsperger, 1959). However, despite the evidence for, at least partially, neurally separate cognitive and affective systems, it is clear that they must interact, in order, for example, for animals to learn about the motivational properties of objects and for emotional experiences to motivate behavior.

If it is accepted that the cognitive and affective systems are (at least, to some extent) empirically separable, investigations into which system is responsible for generating the ERN can be performed. Construction of a situation where an error has no motivational impact (i.e., the error has no negative or detrimental effects and produces no negative affective feelings) would make it be possible to test the two alternatives. If the cognitive system is responsible for the ERN, such an error would still generate an ERN, because the cognitive action monitoring/error detection processes would still be carried out. If, on the other hand, the affective system is responsible for the ERN, such an error might not be expected to lead to an ERN because there would be no motivation to change behavior. While this kind of manipulation might be difficult to achieve because of the negative outcomes strongly associated with errors in everyday life, the experiments here might prove relevant. After the participants have learnt a stimulus-response association, further errors provide no additional information and, therefore, do not motivate a change in behavior.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A similar inquiry might examine different populations to whom errors have a greater or lesser affective impact. If affective processing generates the ERN, it might be expected to be reduced in populations that display weaker affective reactions. Dikman and Allen (2000) investigated the ERN in participants of varying socialization. low socialization is associated with psychopathy and characterized by poor social interactions, antisocial tendencies and lack of responsibility. The results showed that low socialization participants showed a smaller ERN following errors when they were punished for making the error rather than when reward was withheld, again suggesting a link between the ERN and affective processing. Furthermore, studies carried out on other populations with a diminished ERN, such as frontal lobe patients (e.g., Stemmer et al., 2003) and elderly subjects (e.g., Falkenstein et al., 2001; Nieuwenhuis et al., 2002) show they are still aware they have made a mistake.

Another way of testing the alternatives would be to induce the same activity in the affective system (which might also generate a similar feeling) through something other than making a mistake (i.e., using a different input to the affective system). In this case, if the cognitive system is responsible for generating the ERN, we would not expect to observe one because no error detection/action monitoring would have taken place. However, if the affective system produces the ERN, we would expect to see one regardless of what initiated the affective processing. Indeed, Gehring and Willoughby (2002) showed that an ERN-like component (the MFN) was more sensitive to losing money than to accuracy itself, which suggests that, to the extent that the psychological processes underlying the ERN and MFN are similar, the system responsible for the ERN

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

can be activated by negative outcomes (presumably activating the affective system) in the absence of any error processing.

The ERN and Psychological Disorders and Lesions

Aside from the wealth of studies investigating the ERN in normal healthy participants, it has also been studied in a variety of clinical populations and patients with brain damage. Despite the wide psychological and neurophysiological effects of mental disorders, a number of consistent results have been uncovered that give insight into the functional significance of the ERN.

Obsessive-Compulsive Disorder and Tourette's Syndrome

Obsessive-Compulsive disorder (OCD) is characterized by obsessions (intrusive, repetitive thoughts, usually worries, doubts and anxiety) and compulsions (impulsive behaviors performed repetitively in an attempt to relieve the obsessive thoughts). OCD is typically treated with a selective serotonin re-uptake inhibitor (SSRI), frequently coupled with cognitive-behavioral therapy. The effectiveness of SSRIs strongly implicates serotonin in the etiology of OCD.

Several studies into the ERN have been carried out on OCD patients, all of which have demonstrated consistent effects of OCD on the ERN. In the earliest study, Gehring et al. (2000) found that the ERN increased with increasing OCD symptom severity. This finding of an enhanced ERN in OCD patients was replicated by Johannes et al. (2001),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

who also found an increase in negativity on correct trials (an indication of which was also seen Gehring et al.'s results).

A further paper by Hajcak and Simons (2002) found a similar effect in a non-clinical undergraduate population that scored highly on a measure of obsessive-compulsive characteristics (the Obsessive-Compulsive Inventory). The enhanced ERN observed in OCD patients has usually been interpreted as reflecting overactivity of an action monitoring system and relates neatly to OCD symptomatology. The overactive system would continually send error signals, or signals that an action did not have the intended consequence, that could lead to the constant obsessive thoughts and compulsive remedial behaviors.

This result can also be interpreted in the light of the various theories of ERN function. If the ERN is related to error detection, an overactive action monitoring system might result in OCD patients detecting (and probably correcting) more of their errors than controls and erroneously correct already correct responses. While not all of these comparisons have been reported, the behavioral performance of OCD patients and controls in all three papers appears very similar, suggesting that both populations detect errors to the same degree. The conflict-monitoring theory would claim there is an overactive conflict monitoring system that might engage the cognitive control mechanisms to a greater extent. Again this might predict increased correction rates and erroneous corrections. The RL-ERN theory would interpret the enhanced ERN as a stronger reinforcement learning signal. This might predict quicker learning in OCD patients because a stronger

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

learning signal would have a greater impact on the updating of the expectation. Finally, the emotional processing theory would explain the increased ERN in terms of an increased (i.e., worse) emotional evaluation of the event. That is, the event is interpreted (and possibly subjectively experienced) as worse than for controls. This could then focus attention on the event (contributing to obsessions) and motivate repeated remedial behaviors (contributing to compulsions).

Tourette's syndrome shares many similarities with OCD, with which it frequently co-occurs. The ERN in patients suffering from Tourette's syndrome has also been found to be enhanced, similar to that in OCD patients (Johannes et al., 2003). A study by Hajcak et al. (2003) also showed a similar increase in ERN on highly anxious participants. This suggests the increase in ERN observed in OCD and Tourette's syndrome patients may be more related to the anxiety typically experienced with the disorder rather than the disorder itself. Further evidence for the ERN being related to anxiety comes from the finding that following administration of benzodiazepine (anxiolytic medication that is thought to exert its effect as a GABA agonist) the ERN was reduced (Johannes et al., 2001).

The implication of serotonin in the generation of the ERN has implications for the RL-ERN theory, which appeals solely to dopamine as the contributing neurotransmitter. Furthermore, this line of research is not the only one that has implicated the serotonergic system in the production of the ERN. Fallgatter et al. (2004) found that participants with a genetic variation that results in serotonin transporter dysfunction displayed an enhanced

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ERN. This genetic variation is also associated with susceptibility to a number of psychological disorders including OCD and depression.

Schizophrenia

Schizophrenia is a complex, non-unitary disorder that has been the object of several classification attempts. While not all taxonomies agree, there appear, in general, to be three syndromes: reality distortion (characterized by delusions and hallucinations), negative symptoms (characterized by psychomotor poverty and lack of volition) and disorganization (characterized by disjointed thought, inappropriate affect and poor social behavior). Dopaminergic system dysfunction has been implicated in schizophrenia and it is frequently treated with dopamine blockade medication.

Kopp and Rist (1999) were among the first to investigate the ERN in schizophrenic patients. They found that in general the ERN was smaller in patients than controls; however, they noted that the decrease in ERN was considerably greater in paranoid (the reality distortion syndrome) than non-paranoid schizophrenics. They also noted that the error correction behavior in patients was the same as in controls. These results have been replicated by Mathalon et al. (2002), who found a greater decrease in ERN for paranoid than non-paranoid patients but preserved error awareness and post-error slowing. Bates et al. (2002) and Alain et al. (2002) have also demonstrated decreases in the ERN in schizophrenic patients. Furthermore, a study by Bates et al. (2004) showed that ERN amplitude was related to state changes in schizophrenia. Patients were tested off medication during an acute episode and again six weeks following resumption of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

medication. The results showed a highly reduced ERN during medication withdrawal compared with a smaller reduction when they resumed their medication.

The results that demonstrate a decreased ERN but preserved error correction and post-error slowing are problematic for both the error detection and conflict monitoring theories of the ERN. These theories would hold that a disturbance in the error detection or conflict monitoring system would negatively impact higher-level strategic control processes to which they provide input. The RL-ERN might predict that since a reduced learning signal is being sent, schizophrenic patients would learn less quickly than controls. The emotional processing theory might suggest that the errors would not be assigned such emotional significance and, therefore, not capture attention or motivate action as strongly. Again, the normal remedial actions carried out by the schizophrenic patients are problematic for this view.

Parkinson's Disease

The advent of the RL-ERN implicated dopamine in the generation of the ERN; therefore, because Parkinson's disease is associated with dopamine deficiency, studying the ERN in Parkinsonian patients became particularly interesting. Major symptoms of Parkinson's disease are generally psychomotor in nature and frequently include resting tremor and bradykinesia and it is usually treated with medications that increase dopamine levels.

Falkenstein et al. (2002) observed a diminished ERN in medicated Parkinsonian patients on a variety of different tasks. However, Holroyd et al. (2002) found no difference in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ERN between Parkinsonian patients and controls following overnight medication withdrawal. Holroyd et al. offer explanations of these conflicting results in terms of error rates, effects of medication and its withdrawal and response measurement. However, since only two studies have been reported, more systematic research waits to be undertaken to help explain these inconclusive results.

Alzheimer's Disease

One final ERN study that has been carried out on a clinical population concerns Alzheimer's patients. While the aetiology of the disease is poorly understood, it is characterized by brain atrophy and plaque build-up in the brain. Its symptoms include increasing forgetfulness, unclear thought and an inability to complete simple tasks and solve problems.

Mathalon et al. (2003) found that Alzheimer's patients showed a marked decrease in the ERN on a picture-name verification task. However, this result is not surprising given the nature of the disease. A number of systems could be compromised other than the error detection/action monitoring system, such that patients would be less likely to know the correct answer, and more likely to forget their response and even the task instructions.

Aging

There has recently been an upsurge in interest in how cognitive functions change during adulthood (e.g., Park and Gutchess, 2002). This has lead to a number of studies comparing the cognitive abilities of younger and older participants – and ERN research is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

no exception (e.g., Band and Kok, 2000; Falkenstein et al., 2001; Gehring et al., 1995; Mathalon et al., 2003; Nieuwenhuis et al., 2002). These studies have consistently found a reduced ERN in older participants. This cannot be accounted for by a general decrease in ERP amplitude with age because other ERP components were shown to be the same in both younger and older participants (e.g., Stimulus-locked N1, P2 and N2; Nieuwenhuis et al., 2002). Two of the studies (Band and Kok, 2000; Falkenstein et al., 2001) also report older subjects demonstrate error correction and post-error slowing despite the decrease in the ERN. Again, these results highlight the problem of normal (or, at least, relatively spared) error correction and strategic changes accompanying a major reduction in ERN amplitude.

Lesions

Owing to its central location in the brain, it is relatively rare to find patients with selective ACC damage. However, patients suffering from a stroke affecting the anterior communicating artery or the pericallosal branch of the anterior cerebral artery may suffer from focal ACC damage. It was on these types of patients that Swick and Turken (2002) and Stemmer et al. (2003) investigated the ERN. Both studies demonstrated a decrease in the ERN for patients compared with controls although some patients retained awareness of their errors and post-error slowing. Swick and Turken (2002) also found that their patient made more errors on incongruent trials of a Stroop task (i.e., they were more sensitive to distracting stimuli), which suggests an attentional impairment.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Two further studies have investigated the ERN following damage to fronto-lateral brain regions. Gehring and Knight (2000) showed that lesions to the dorso-lateral prefrontal cortex (DLPFC) increased ERN-like activity on correct trials, such that there was no difference between error and correct trials. Ullsperger et al. (2002) found that the ERN was abolished in patients with fronto-lateral damage, but not in those with orbito-frontal or temporal damage. Although the effects of the lesions on the ERP were different in these studies, they both suggest the ACC and DLPFC are functionally connected and part of a broader system.

The DLPFC has been hypothesized to have a number of possible functions. One is that it holds information in memory, such as the task requirements and response representations. The other is that it exerts top-down control on the action-monitoring system by, for example, focusing attention. Damage to a structure serving either function might be likely to disrupt an action monitoring system. A system that could not hold response representations in memory would not be able to compare them and detect a mismatch between them. A system that could not implement top-down control would not be able to make strategic changes to perform tasks optimally.

Research into the Anterior Cingulate Cortex

If the current hypothesis that the ERN is generated by the ACC is correct, other clues to its functional significance might be uncovered by examining what has been learnt about ACC function using different methodological techniques and investigating different psychological processes. Studying the behavior of patients with damage to the ACC can

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

give insight into its function, as can neuroimaging techniques such as fMRI and Positron Emission Tomography (PET) which, instead of measuring electrical brain activity, measure regional cerebral blood flow (rCBF). Those areas of the brain that are more active require more metabolic energy; therefore, blood flow to those parts of the brain increases. While it is clear that electrophysiology and fMRI and PET measure different physical processes, it is hoped that, at least to some extent, they may index the same psychological process. The relationship between these differing neuroimaging techniques is far from resolved; however, a number of studies suggest that source localization from electrophysiological data and rCBF during the same tasks are very similar, which suggest they may indeed be measuring brain activity during the same psychological process (e.g., Grimm et al., 1998; Sommer et al., 2003; but see Vitacco et al., 2002). However, it is not claimed here that the ERN found in electrophysiological experiments is functionally equivalent to ACC activation observed using fMRI and PET, merely that some insight into ACC function and, therefore, the functional significance of the ERN might be gained through other methodologies.

The ACC

Some of the earliest studies to give insight into the functions of the ACC investigated patients with medial-frontal lesions. For example, Barris and Schuman (1953) described one such patient as showing apathy, akinesia and an indifference towards pain. They also reported abnormalities in the patient's heart rate, respiration and temperature. They concluded that the ACC was involved in volition and autonomic control – both of which appear in later research. Indeed, Oschner et al. (2001) reported a patient with ACC

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

damage who was unable to initiate novel motor behavior and Critchley et al. (2003) demonstrated ACC activity related to heart rate modulation that was independent of cognitive or motivational processing, suggesting that the ACC is involved in autonomic control. It was also found that patients with ACC damage had abnormal heart rate control and decreased autonomic arousal. This can be related to ERN research that shows heart rate modulation following error feedback in a task during which the ERN is typically elicited (Crone et al., 2003).

A different line of research involves electrical stimulation of the ACC, which has resulted in a number of effects. Talairach et al. (1973) elicited involuntary sequences of movements, such as touching, rubbing or stretching on sustained stimulation of the ACC. Bancaud and Talairach (1992) found they could elicit different types of emotion depending on what part of the ACC they stimulated. Burns and Wyss (1985) measured changes in various autonomic functions such as blood pressure, respiration and pupillary dilation with ACC stimulation. This suggests the ACC is involved in emotional processing, autonomic control and closely coupled with motor processing.

In two reviews of ACC structure, function and functional connectivity, Vogt and his colleagues (Devinsky and Morrell, 1995; Vogt et al., 1992) suggest that the ACC is an interface between cognition and motivation. While the exact nature of the interface is not specified, evidence for both structural and functional subdivisions within the ACC is cited, with the rostral portion of the ACC specialized for affective processing and the more caudal region involved in cognitive processing. More recent cytoarchitectural

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

research has provided further evidence for structural subdivisions within the ACC (Vogt et al., 2003). Neuropsychological and neuroimaging research has also provided evidence for functional subdivisions within the ACC. For example, Swick and Jovanovic (2002) found that lesions to different areas of the ACC result in different behavioral performance and a meta-analysis of neuroimaging studies that employed cognitive and emotional tasks showed that the cognitive tasks activated more caudal areas of the ACC, whereas emotional tasks tended to active more rostral areas (Bush et al., 2000).

Owing to its close connections with other limbic structures, the ACC is frequently regarded as being an evolutionarily comparatively old region of the cortex. However, this view has recently been challenged by the discovery of a new class of spindle-shaped neurons that appear to be found only in humans and the great apes (Allman et al., 2001). They are characterized by two large dendrites, and apical one projecting up and a basal one projecting downwards (giving them their spindle shape). Although not much is known about these cells, they are hypothesized to have widespread cortical connections and be involved in integrating a wide variety of information enabling increased problem solving abilities. This type of function would be consistent with the putative integrative nature of ACC function.

The ACC and Action Monitoring

Because electrophysiological methods have implicated the ACC in action monitoring, its involvement in this function has been investigated further using other neuroimaging techniques. And, indeed, ACC activity has been observed following errors in a number

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of different tasks (e.g., Braver et al., 2001; Fassbender et al., 2004; Garavan at al., 2002; Kiehl et al., 2000; Menon et al., 2001; Ullsperger and von Cramon, 2003). Furthermore, while precise localization is difficult because the areas of activity are frequently quite large and there is no universally accepted anatomical division of the ACC, most of these studies report error related activity in the rostral ACC. However, a meta-analysis of neuroimaging studies by Ridderinkhof et al. (2004) showed overlapping regions of in dorsal ACC, the pre-supplementary motor area (pre-SMA) and the supplementary motor area (SMA) activated by errors, conflict, uncertainty and negative feedback. A further study investigated ACC activity on trials where participants either recognized their errors immediately or they were presented with error feedback (Holroyd et al., 2004); this parallel the research into the response ERN and feedback ERN. The results showed that the same area of (dorsal) ACC was activated on both types of trial. While it might be unwarranted to assume ACC activity measured by fMRI is equivalent to the ERN, this result is certainly suggestive.

Owing to the ACC's implication in monitoring response conflict (e.g., Botvinick et al., 2001), several studies have tried to dissociate ACC activity related to error processing from that related to conflict monitoring (e.g., Fassbender et al., 2004; Garavan et al., 2002). These studies have typically shown greater ACC activity following errors in more rostral regions, whereas conflict related activity appears located in more dorsal and superior areas towards the pre-SMA and SMA, suggesting a functional dissociation of conflict and error processing within the medial-frontal cortex. This dissociation has also been proposed on the basis of neuropsychological research. Swick and Turken (2002)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

observed a diminished ERN in a patient with rostral ACC damage, but a larger N450, which is thought to reflect conflict related processing. A further connection to the ERN literature comes from research into schizophrenic patients. Laurens et al. (2003) found decreased rostral ACC activity in schizophrenic patients than controls. Again, this parallels the decreased ERN observed in schizophrenic patients.

The ACC and Attention

ACC activity has frequently been observed in a number of tasks commonly thought to engage attentional processes, such as the Stroop and flanker tasks (e.g., Casey et al., 2000; Peterson et al., 1999). Inspired by the conflict monitoring research, these and similar studies (e.g., Bush et al., 1998; Erickson et al., 2004; Milham et al., 2001; Milham et al., 2003) typically claim the ACC is recruited to mediate and resolve response conflict. An interesting finding from some of these studies (e.g., Bush et al., 1998; Milham at al 2003; Erickson et al., 2004) is that ACC activity decreases over the course of the task. The interpretation has been that the ACC is only involved early in the task in engaging other areas that provide ongoing attentional control.

However, there is also evidence that the ACC is active during covert shifts of attention and not during response related processes, such as response selection or preparation (Benedict et al., 2002). Furthermore, Paus et al. (1997) found that ACC activity during a vigilance task covaried with thalamic activity rather than other frontal activity. This was interpreted as providing evidence that the ACC activity was more related to the arousal or alertness component of attention than response related processes. Neuropsychological

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

studies have also implicated the ACC in attentional functions. For example, both Swick and Jovanovic (2002) and Danckert et al. (2000) observed that ACC damage resulted in altered patterns of interference from irrelevant information, suggesting the attention system was somehow compromised.

The ACC and Emotion and Pain

The ACC is generally considered part of the limbic system and, as such, has a hypothesized role in emotional processing. While the specific nature of the emotional processing remains unspecified, many studies have demonstrated ACC activity during emotional tasks (e.g., Barrett et al., 2004; Lane et al., 1998). In fact, Phan et al. (2002) carried out an extensive review and meta-analysis of many neuroimaging studies of emotion and showed that the ACC was consistently activated during a wide variety of emotional tasks. A review concentrating more specifically on the medial-frontal cortex (Sewards and Sewards, 2003) suggested that representations of motivational drives, such as fear, thirst and hunger were located in the more rostral region of the medial-frontal cortex, again consistent with its hypothesized role in emotional processing. Furthermore, following rTMS of the medial-frontal cortex, discrimination of angry (but not happy) faces was compromised (Harmer et al., 2001).

ACC dysfunction has also been implicated in a number of psychological disorders (those that are usually understood in terms of abnormal emotional processing). For example, ACC dysfunction has been linked to schizophrenia (e.g., Szesko et al., 2000; Tamminga et al., 2000), obsessive-compulsive disorder (e.g., Levin and Duchowny, 1991),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

depression (e.g., Davidson et al., 1999; Ebert and Ebmeier, 1996), autism (e.g., Burack et al., 2001) and psychopathy (e.g., Kiehl et al., 2001). Furthermore, Cohen et al. (2001) found that following cingulotomy, patients displayed a reduction in tension, anger and anxiety.

However, there is evidence that it is not just emotional processing that activates the ACC. Fichenholtz et al. (2004) found that both emotional stimuli and non-emotional, but motivationally significant stimuli, activated the ACC. This suggests the ACC is not responsive to emotional stimuli per se, but motivationally significant stimuli in general. Consistent with these results (and paralleling some of the findings of research into the ACC and its attentional function), Phan et al. (2003) discovered that ACC activity decreases on the second presentation of emotionally salient pictures. An interpretation might be that emotional stimuli that are usually motivationally significant and, therefore, activate the ACC, lose their significance if they have already been seen, and fail to activate the ACC on subsequent presentation.

Pain is closely associated with emotion because the experience of pain usually results in a negative emotional or evaluative processing and a negative affective state. It is not surprising, therefore, that pain processing is strongly associated with the ACC. Neuroimaging studies (e.g., Lorenz et al., 2003; Ringler et al., 2003) have shown that ACC activity, when people are subjected to painful experiences, varies with pain intensity. However, it has also been shown that the ACC does not code the sensory component of pain but rather the evaluation of it or its motivational significance. For

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

example, under hypnotic analgesia, ACC activity varied with the perceived unpleasantness of the pain (Faymonville et al., 2003; Rainville et al., 1997). In a similar vein, Kropotov et al. (1997) recorded intracranial potentials in the ACC of two OCD patients, one of whom was responsive to hypnotic suggestion. The results showed that only in the responsive patient did ACC activity (caused by a painful stimulus) decrease during hypnotic analgesia. A further study by Wager et al. (2004) investigated the neural mechanism of placebo anesthesia, in which reduction in pain perception was associated with decreased ACC activity. The complementary results have also been observed, those of pain affect without pain sensation. Ploner et al. (1999) studied a patient with somatosensory cortex damage who could not feel pain yet experienced its negative affect.

Experiencing pain itself, however, is not the only way to activate the ACC. Singer et al. (2004) demonstrated watching someone else experience a painful stimulus activated the ACC; moreover, the degree of ACC activation correlated with empathy. Osaka et al. (2004) showed that hearing words expressing pain could also activate the ACC and Eisenberger et al. (2003) demonstrated that social exclusion activated an ACC region similar to that activated by physical pain. Results such as these suggest the ACC is involved in processing negative evaluative feelings, or feelings of unpleasantness, that are associated with physical pain itself, but can also be elicited by related stimuli.

The ACC in Animals

Single cell recordings and focal lesions in animals can elucidate brain function in a way that is typically unavailable in humans. One of the advantages of animal research is that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

it allows precise localization of electrode implantation and lesions. There are, however, a number of disadvantages. Inferences from animal brain organization to that of humans must be made. Also, while investigation of individual neuronal responses can give insight into brain function, it overlooks the functional processing of populations of neurons as a whole. Nevertheless, some animal research is germane to a discussion of ACC function.

One of the earliest papers that investigated ACC function in animals found a number of different types of cell, one of which was responsive to errors and the omission of an expected reward (Niki and Watanabe, 1979). These cells have since been implicated in a potential neural mechanism for the generation of the ERN. This and subsequent research (e.g., Ito et al., 2003; Shidara and Richmond, 2002) has suggested the ACC contains neurons that code for the consequences of events and their anticipation.

However, other research has suggested a slightly more sophisticated view of ACC neurons. For example, Shima and Tanji (1998) found cells that only fired following a reduced reward when that reduction motivated a change in behavior. This suggests motivational properties of events might have an important role to play on top of simple reward expectancy. Furthermore, ACC neurons have been implicated in learning. Gabriel et al. (2002) and Cardinal et al. (2002) have both suggested that cells in the ACC are involved in learning goal-directed behavior based on specific stimuli and their association with particular outcomes. In an interesting parallel to the neuroimaging

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

research, Gemba et al. (1986) found that intracranial potentials generated by the ACC in monkeys were only seen after errors early in learning.

Other studies have discovered more parallels with human neuroimaging research. Koyama et al. (1998) found neurons in the monkey ACC that were responsive to pain and pain anticipation. Also, Walton et al. (2003) found that rats with ACC lesions engaged in high effort behaviors much less than they had done prior to the lesions, again, suggesting a link to motivational processes.

Testing Theories of ERN Function

There are a number of consistent themes running through research into the ERN and ACC and parallels can be drawn across different methodological approaches. Much of the evidence suggests the ERN might reflect a motivational function, which is carried out in the ACC. The experiments presented here, the results of which are consistent with a motivational function of the ERN, were designed specifically to test predictions made by the RL-ERN theory and their results can be interpreted regarding that theory. However, they also have relevance to other theories of ERN function and can be interpreted with respect to them. Finally, the results can be put in a broader picture of ACC function in an attempt to synthesize a wider body of knowledge.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter II

Experiment 1

Experiment 1a

Design and Rationale

Holroyd and Coles's RL-ERN model is a novel conceptualization of the functional significance of the ERN. Therefore, it makes new predictions that can be tested using new experimental designs. This not only examines the RL-ERN but also gives us further insight into the functional significance of the ERN in general.

One of the key predictions of the RL-ERN model is that the amplitude of the ERN will be greater the worse an outcome is from expected. Using various probability learning tasks, this prediction is tested explicitly here. In Experiment 1, participants learnt a probabilistic relationship between a stimulus and an outcome (winning or losing money). Participants' expectations of correct or incorrect feedback changed as they learnt the relationship, such that (at least for the non-random associations) they would increasingly expect correct feedback and error feedback would become increasingly unexpected. The degree of participants' expectation was inferred from subjective confidence ratings and from how long they had been learning the relationship between the stimulus and the outcome. If the prediction derived from the RL-ERN model is correct, the ERN will be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

largest following error feedback that participants are not expecting, that is when they have high confidence that their response is correct or later in learning, when they have had time to learn the relationship and build up expectations. If, however, this pattern is not observed, the RL-ERN model's predictions will not be upheld, which will have implications for the RL-ERN model specifically and the functional significance of the ERN in general.

Methods

Participants

There were 8 participants (6 males and 2 females) aged between 19 and 25. All were right-handed, had normal or corrected-to-normal vision and were students at the University of Michigan. They were paid $8 per hour plus bonus money accumulated over the course of the experiment (described below).

Procedure

Participants were instructed to learn the probabilistic relationship between stimuli and outcome values associated with them. The stimuli were random 6-sided polygons taken from the collection of Attneave shapes (Attneave, 1957). The outcome values were -20, -2, +2 and +20, representing a loss or a gain of that amount of money in US cents. There were 6 conditions: 3 conditions with the outcome values -2, +2 and +20 strongly associated with stimuli (85% of the time), 2 conditions with values -2 and +20 weakly associated with stimuli (55%) and a random association (25%) condition. The remaining association probability was divided equally between the other outcome values. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

probability mappings between the stimuli and outcomes used in Experiment 1a are shown in Table 2.1 with some sample Attneave shapes.

Experiment 1a		Outcome Value				Sample Attneave Shape
		-20	**-2**	**+2**	**+20**	
Association Strength (%)	Strong	5	5	5	**85**	
		5	5	**85**	5	
		5	**85**	5	5	
	Weak	15	15	15	**55**	
		15	**55**	15	15	
	Random	25	25	25	25	

Table 2.1 Association strengths for each outcome value in Experiment 1a and example Attneave shapes.

Participants were seated in a sound treated, dimly lit testing chamber about 60cm from the monitor, where the stimuli subtended approximately two degrees of visual angle. They were first shown a shape and, in response, had to select which value they thought was associated with it by pressing one of four buttons on a response pad (on which four fingers of their right hand rested). Following this response, the participants were asked to rate their confidence in their value selection, again by pressing one of the 4 buttons, which now represented different degrees of confidence (low, 0-25%; medium-low, 25-50%; medium-high, 50-75%; high, 75-100%). Next, a symbol was presented indicating whether their value selection was correct or incorrect. The symbol was @ or #, the meaning of which (correct or error) was counterbalanced across participants. Finally, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

actual value associated with the shape on that particular trial was presented. This was either 02 or 20 presented on either a blue or yellow background to signify either a win or a loss. Again, the meaning of the color (win or loss) was counterbalanced across participants. Participants actually won or lost the amount of money presented on a trial-by-trial basis; if they accrued a net gain over the course of the block this amount was added to their payment. However, if they accrued a net loss, no money was subtracted. Figure 2.1 presents a schematic representation of the task.

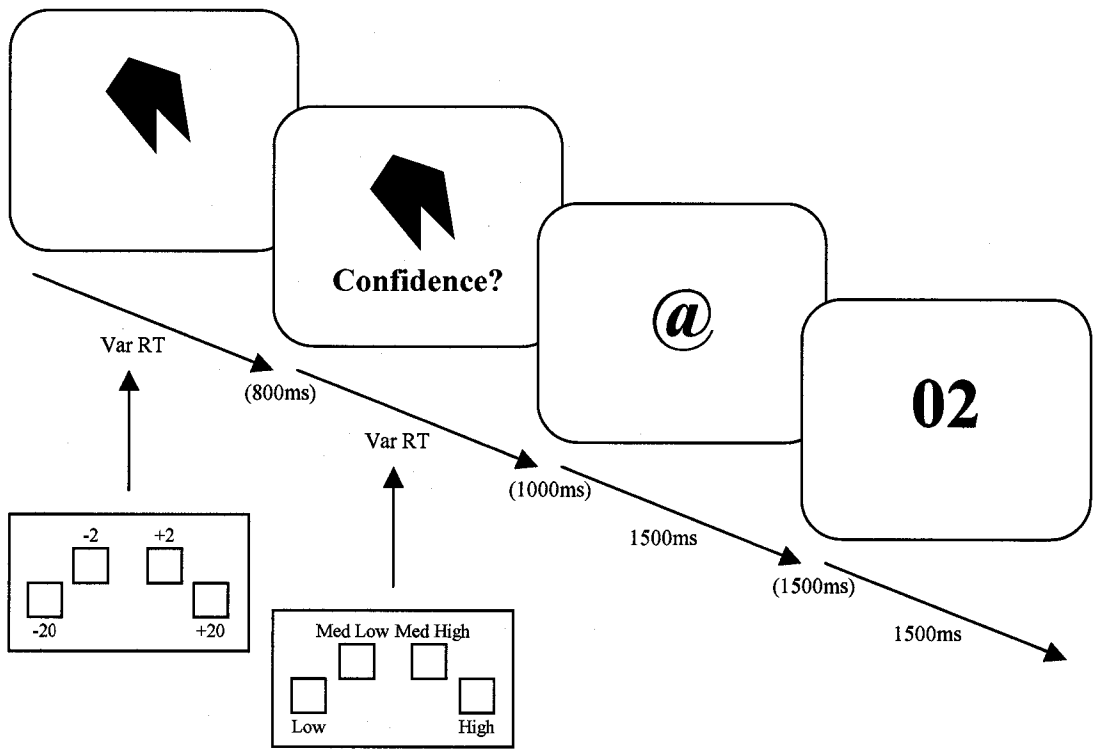


Figure 2.1 A schematic representation of Experiment 1.

Participants completed 8 blocks (each of 90 trials) of the associative learning task. During each block, 3 different shapes were presented. Therefore, each participant completed 4 blocks of each of the 6 probability mappings. Each shape was randomly assigned to one of the 6 probability mappings, with the constraint that no two of the same

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mapping could be presented in any individual block. Each block was divided into 3 parts (early, middle and late, each of 30 trials) and after each part the participant was required to estimate the association between each shape and the four outcome values (i.e., estimate the probability mapping).

Electrophysiological Methods

The EEG was recorded from 37 Ag/AgCl electrodes (Falk-Minow, Germany) positioned at sites Nasion, Fp1, Fp2, AFz, F9, F7, F3, Fz, F4, F8, F10, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P9, P7, P3, Pz, P4, P8, P10, O1, Oz, O2 and Inion according to the International 10-20 System (Jasper, 1958). The left mastoid was used as a reference, with an average mastoid reference derived off-line using right mastoid data. Horizontal and vertical electrooculograms (EOG) were recorded from electrodes placed outside the outer canthus of each eye and above and below both eyes, respectively. EEG and EOG were amplified by SYNAMPS high impedance amplifiers (Neuroscan, Inc.) and low-pass filtered on-line at 50 Hz. The data were digitized at 1000 Hz and subsequently reduced to a sampling rate of 200 Hz. Ocular artifacts were removed from the EEG data using the procedure in Gratton et al. (1983). The average value in the 100 ms preceding stimulus presentation (i.e., -100 – 0 ms as shown on the waveform plots) was used as the baseline.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Results

Behavioral Results

The behavioral results of Experiment 1a generally show what might have been expected. Both accuracy and confidence increased over the block and were higher for the stronger associations. This suggests the participants were learning the associations and becoming more confident in their value selections as learning progressed. Figure 2.2 illustrates the increase in accuracy over the three parts of the block. A one-way repeated measures ANOVA showed a significant increase in accuracy over the three parts ($F(2, 14) = 11.75$, $MSE = 0.0066$, $p = 0.001$; all proportions have been subjected to the arcsine transform and Greenhouse-Geisser adjustments have been made where appropriate).

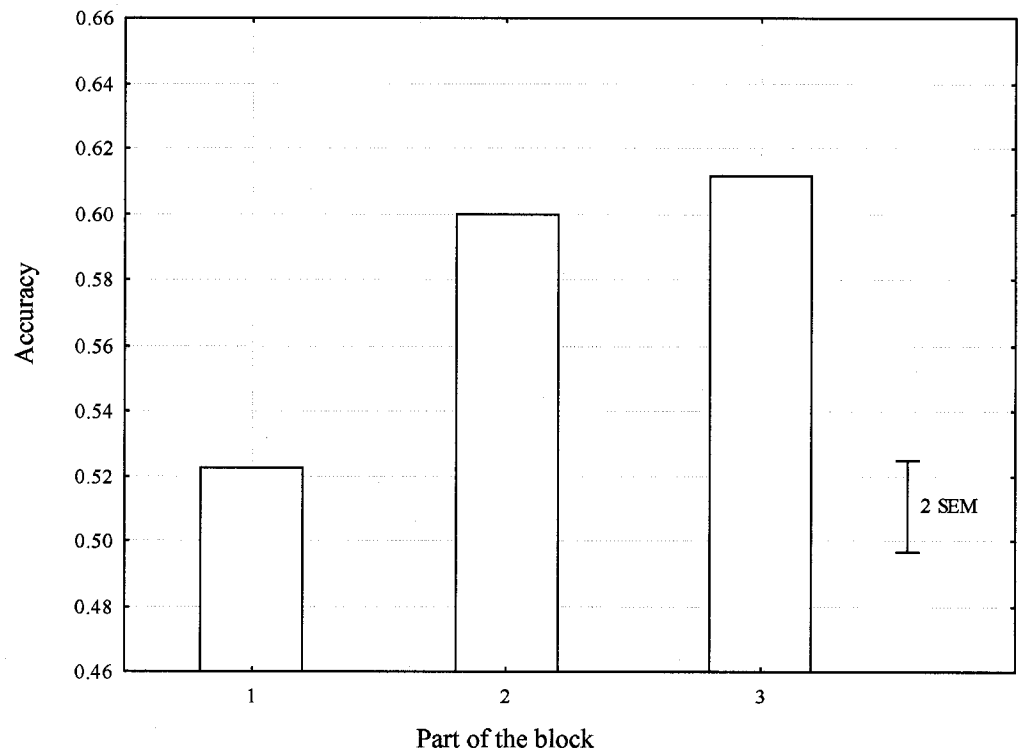


Figure 2.2 Experiment 1a change in accuracy over the block. Accuracy increased over the course of the block. The error bar in this and all graphs represents 2 standard errors of the mean, calculated from the corresponding MSE term in the ANOVA.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A subsequent Tukey HSD test revealed there to be a significant difference between part 1 and part 2 ($p = 0.0047$) and part 1 and part 3 ($p = 0.0016$) but not between parts 2 and 3 ($p = 0.83$). The difference in accuracy for the different association strengths was also significant ($F(5, 35) = 82.15$, $MSE = 0.022$, $p < 0.0001$). Figure 2.3 shows this effect was caused by the differences in accuracy between the random, 55% and 85% conditions. A post-hoc Tukey HSD test showed mean accuracy in all the conditions to be different from each other ($p < 0.0005$). There were no significant differences in accuracy within the same probability associations. This illustrates participants were more accurate for stronger association strengths.

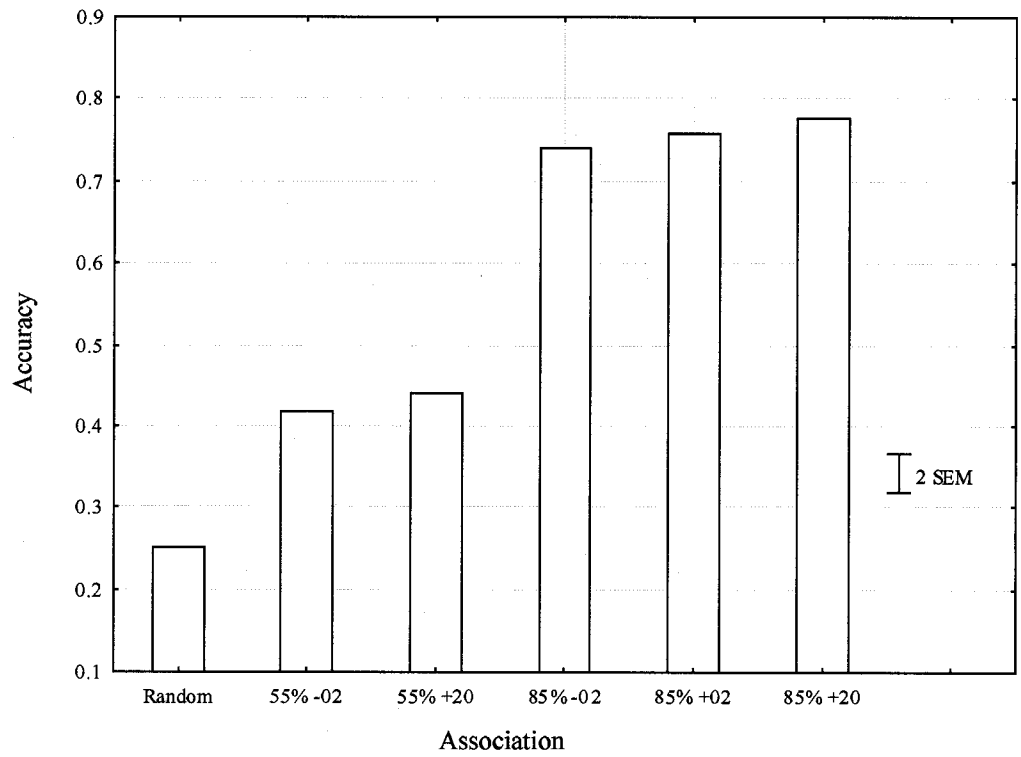


Figure 2.3 Experiment 1a accuracy for the different associations.

The results also showed that participants were significantly more accurate when they were more confident ($F(3, 21) = 43.02$, $MSE = 0.022$, $p < 0.0001$). Figure 2.4 shows

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

accuracy for the four confidence ratings. A Tukey HSD test revealed the accuracy at each confidence level was different from that at the adjacent level (low vs. medium-low: $p = 0.009$, medium-low vs. medium-high: $p = 0.004$, medium-high vs. high: $p = 0.026$).

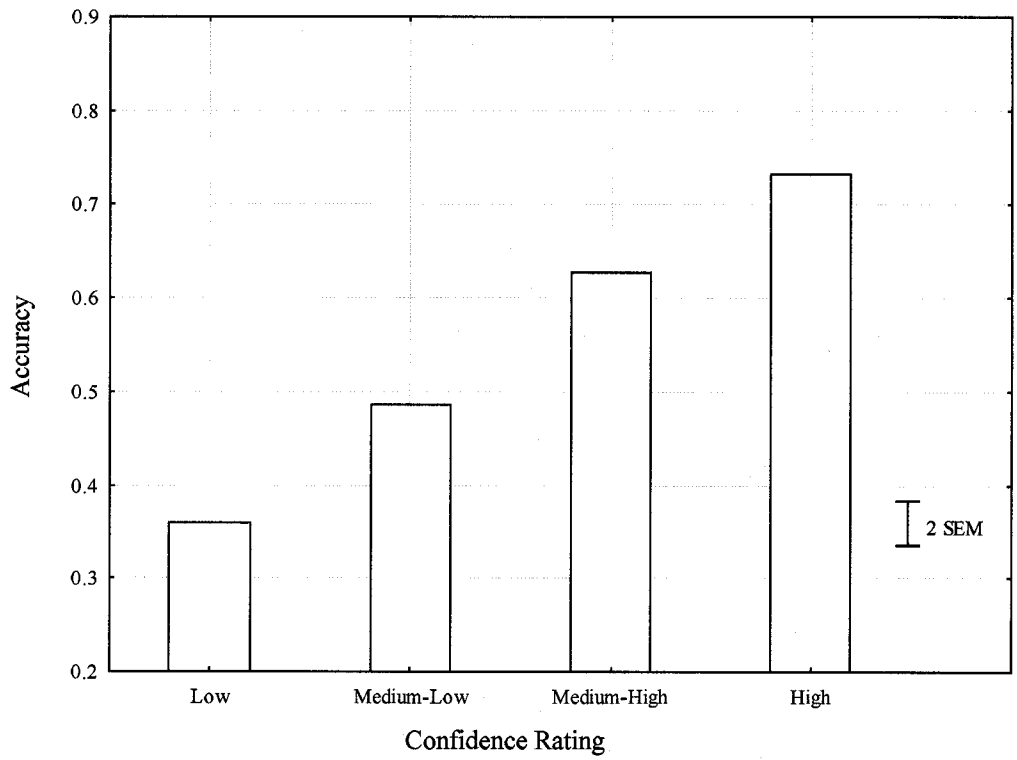


Figure 2.4 Experiment 1a accuracy for the different confidence ratings.

Similar effects were also seen for confidence as well as accuracy. Confidence increased over the three parts ($F(2,14) = 13.07$, $MSE = 0.0827$, $p = 0.0035$) and increased with increasing association strength ($F(5, 35) = 24.53$, $MSE = 0.099$, $p < 0.0001$). The finding that confidence increases over time suggests that these variables are related and that both can be used to estimate participants' expectations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Electrophysiological results

Potentially, both a response and feedback ERN could be observed in this experiment. However, a response ERN would be unlikely because no stimulus is perfectly associated with only one outcome, so participants could never be sure of the accuracy of their response before the accuracy feedback was presented. Also, because participants gained and lost money, there is the possibility that MFN activity would be observed following feedback indicating a monetary loss. In fact, no response ERN was observed in any of the conditions, nor was any MFN activity seen following monetary losses. Therefore, most of the analyses presented here examine the feedback ERN

Figure 2.5 compares the overall waveforms following error and correct feedback averaged across all conditions and all participants at electrode FCz (where the ERN is frequently maximal). An ERN is observed that peaks at 236 ms following error feedback (all waveforms are shown filtered with Chebyshev type II low-pass filter with a half-amplitude cut off at 12 Hz). While the error waveform shows more negativity than does the correct, this difference is not significant ($F(1, 7) = 3.12$, $MSE = 3.52$, $p = 0.12$). For statistical purposes, comparisons are made between mean amplitude measurements taken 50 ms either side of the negative peak of the overall ERN at electrode FCz unless otherwise stated (236 ms for Experiment 1a, 234 ms for Experiment 1b). The topographic map of the ERN (the error-correct difference) at 236 ms shows a medial-frontal distribution with the difference maximal at electrode FCz.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

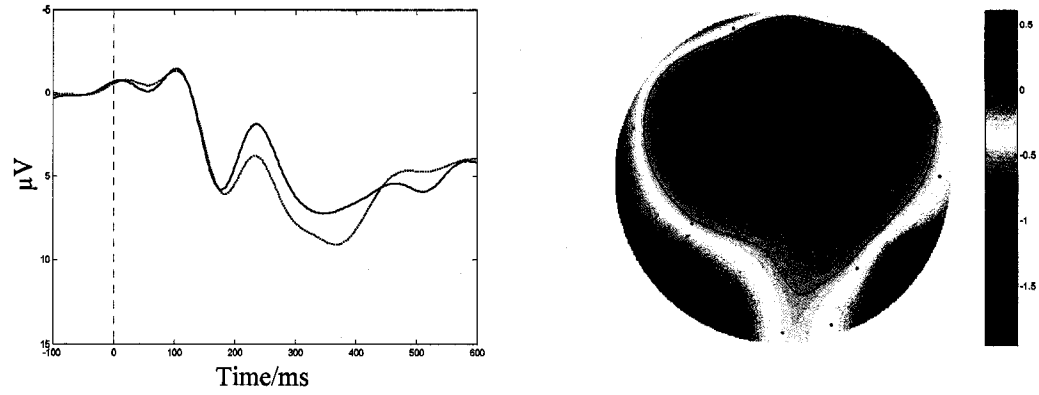


Figure 2.5 Experiment 1a ERN averaged over all conditions. The left panel shows the waveforms following correct and error feedback. In this and all other graphs the solid line is the error waveform and the dotted line is the correct waveform (unless otherwise stated). The topographic map in the right panel shows the scalp distribution of the difference between the error and correct waveforms at 236 ms (units are μV).

Inspecting the waveforms on error and correct trials over the course of the block illustrates why, when averaged over the three parts, the difference between them was not significant. Figure 2.6 shows that the ERN was much greater in the first part than the second and third. This was reflected by a significant interaction in a 3 (part) x 2 (accuracy) ANOVA ($F(2, 14) = 6.56$, $MSE = 1.21$, $p = 0.017$). The ERN (the difference between the error and correct waveforms) was only significant in part 1 ($F(1, 7) = 13.93$, $MSE = 3.07$, $p = 0.0073$); and was significantly larger in part 1 than parts 2 or 3 (part 1 vs. part 2: $F(1, 7) = 9.09$, $p = 0.020$; part 1 vs. part 3: $F(1, 7) = 18.63$, $p = 0.0035$). There was no difference in the ERN between parts 2 and 3 ($F(1,7) = 0.07$, $p = 0.80$).

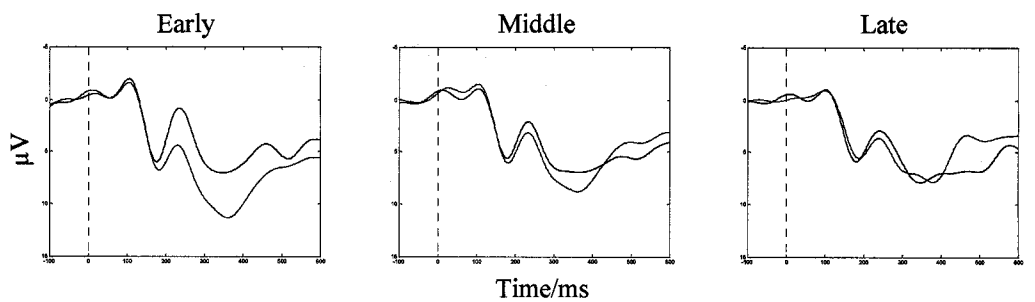


Figure 2.6 Experiment 1a change in ERN over the block.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The ERN was also modulated by confidence. Figure 2.7 shows the error and correct waveforms when grouped by confidence rating. The ERN appeared to decrease as participants' confidence in their choices increased. The interaction in a 4 (confidence) by 2 (accuracy) ANOVA, while not significant ($F(3, 21) = 2.76$, $MSE = 4.43$, $p = 0.096$), is suggestive that the ERN might change with confidence. Looking more closely at the waveforms for individual confidence ratings the results show that the difference between the error and correct waveforms was significant for the two lowest confidence ratings (low confidence: $F(1, 7) = 7.68$, $MSE = 9.55$, $p = 0.028$; medium-low confidence: $F(1, 7) = 12.56$, $MSE = 3.46$, $p = 0.0094$) but not the highest two (medium-high confidence: $F(1, 7) = 0.85$, $p = 0.39$; high confidence: $F(1, 7) = 0.06$, $p = 0.81$).

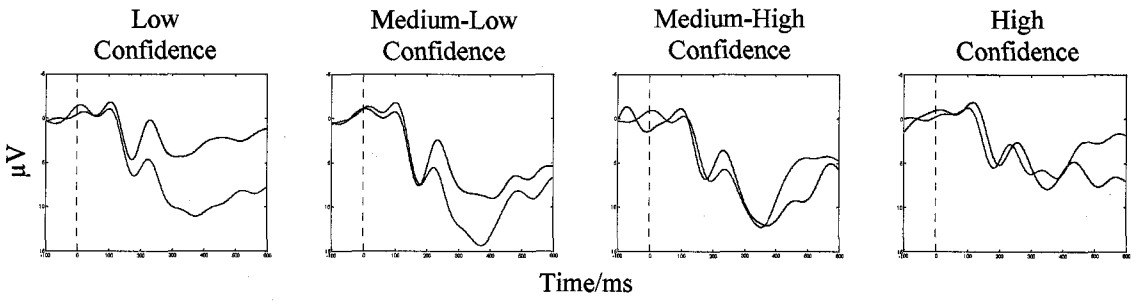


Figure 2.7 Experiment 1a ERN for the four confidence ratings.

These results suggest the four confidence rating groups can be collapsed into two (low and high), by combining the two lowest and two highest ratings. When this was done, a 2 (confidence) x 2 (accuracy) repeated measures ANOVA showed a significant interaction ($F(1, 7) = 6.95$, $MSE = 2.05$, $p = 0.034$), reflecting a difference in the size of the ERN between the low and high confidence groups (Figure 2.8). This analysis also showed there was a significant difference between the error and correct waveforms in the low confidence grouping ($F(1, 7) = 22.83$, $MSE = 2.20$, $p = 0.002$) but not in the high confidence grouping ($p = 0.56$). Figure 2.8 also shows the scalp distributions for low and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

high confidence ERNs (at their respective peaks, 234 and 246 ms). These maps show a much greater medial-frontal negativity for low confidence ERNs than high confidence ones. According to the RL-ERN model, errors following high confidence judgments should be the most unexpected and, therefore, result in the largest ERN.

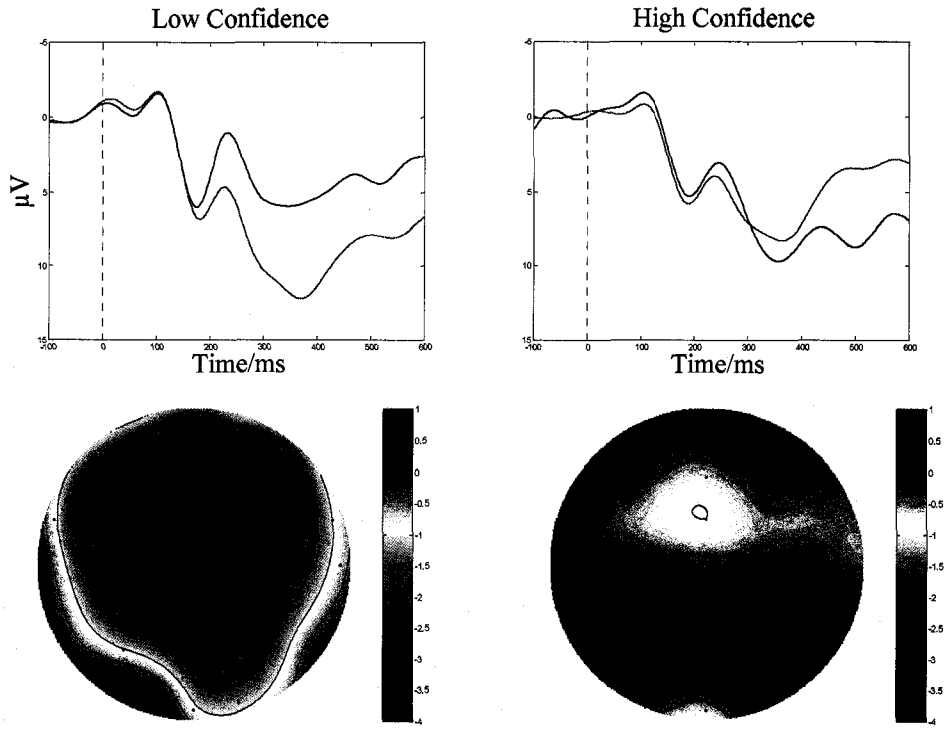


Figure 2.8 Experiment 1a ERN for grouped confidence ratings. The lowest two confidence ratings have been to form the low group and the highest two have been grouped to form the high group. The topographic maps show the scalp distribution at the peak of the negativity in the error waveform (units are μV).

One question that arises is whether the ERN is actually modulated by position in the block or by subjective confidence in the choice. Because confidence is low at the beginning of the block and increases over time, position in the block and confidence are correlated. However, they are distinct; while it might not be possible (in this experimental paradigm) to have high confidence early in learning, it is possible to have low confidence in a value selection when the association strength had been learnt (the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

random condition). A practical problem, however, is that there will be few low confidence trials later in the block, and few high confidence trials early in the block, resulting in fewer waveforms to average. Furthermore, these trials will be in different proportions for the different association strengths. The late low confidence trials will be more likely to come from the random and 55% association strengths than the 85% condition and the early high confidence trials will be more likely to come from the 85% association strength condition. Nevertheless, Figure 2.9 suggests that the ERN depends more on the subjective confidence of the choice rather than the actual position of the error in the block.

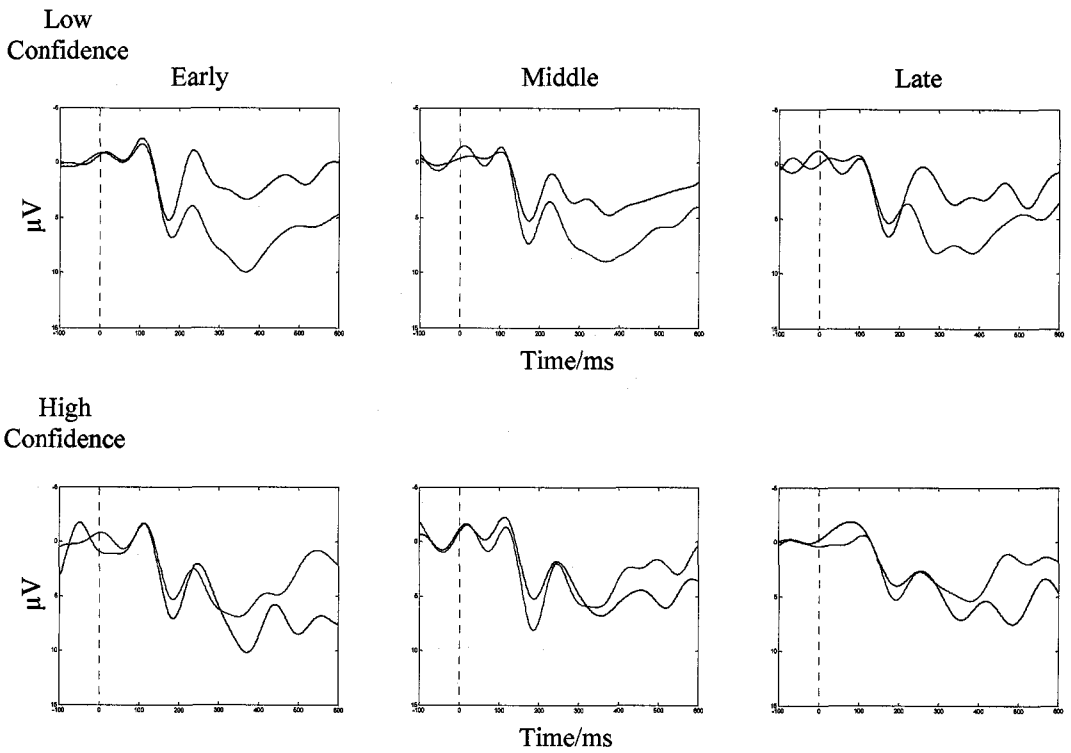


Figure 2.9 Experiment 1a change in ERN over the block for the low and high confidence ratings.

None of the differences between error and correct waveforms were significant in the high confidence conditions. In the low confidence conditions, the difference between the error and correct waveforms in the early part of the block was significant ($F(1, 7) = 33.83$,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MSE = 3.097, p = 0.0007), in the middle part it approached significance ($F(1, 7) = 4.55$, MSE = 7.52, p = 0.070) and in the late part it was not significant (p = 0.18). Furthermore, the ERN was larger in the low confidence grouping than the high confidence grouping for both the early and middle parts (early: $F(1, 7) = 6.39$, MSE = 6.02, p = 0.039; middle: $F(1, 7) = 5.66$, MSE = 2.27, p = 0.049).

An alternative way of conceptualizing errors of differing confidence ratings is to divide errors into two categories – ignorance errors and violation errors. An ignorance error is defined as an error that is made when the participant has not learnt the associative relationship between the stimulus and the outcome. That is, the participant is presented with error feedback after selecting any of the values followed by the lowest confidence rating. A violation error is defined as an error that is made when the participant has some knowledge of the stimulus-outcome relationship. That is, the participant is presented with error feedback after selecting the probabilistically most likely outcome and choosing one of the three highest confidence ratings. Figure 2.10 shows the waveforms for ignorance and violation errors compared with analogous trials when participants were presented with correct feedback. The scalp distribution of the ERN at the peak negativity of the error waveform (within a 200-300 ms window) again shows a much greater fronto-central negativity for ignorance errors.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

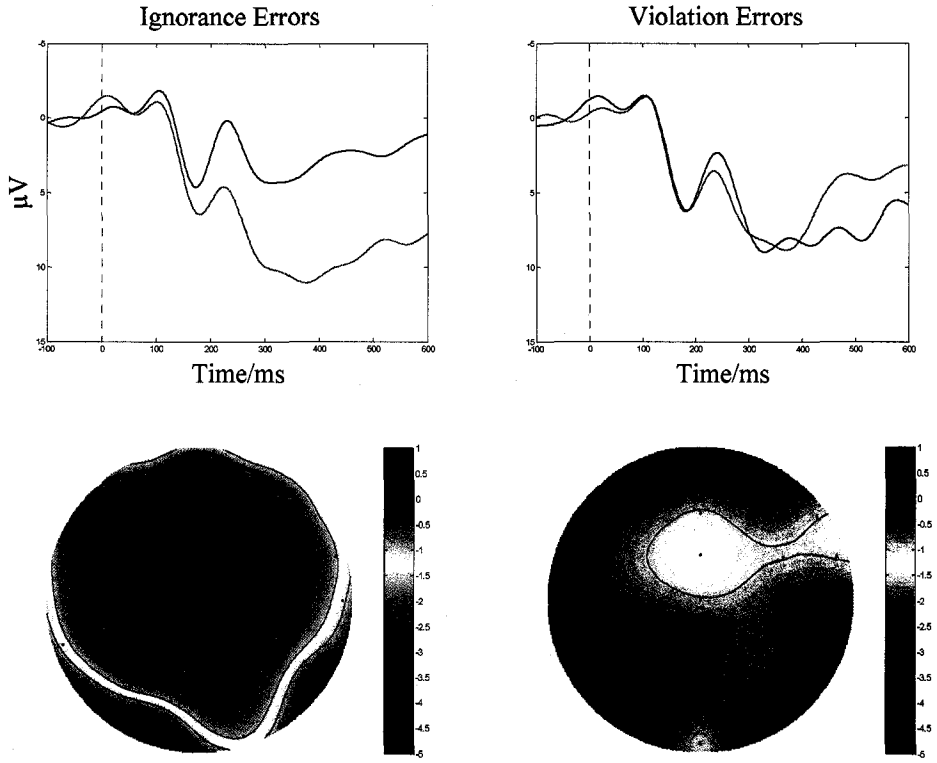


Figure 2.10 Experiment 1a ignorance and violation errors. The topographic maps show scalp distributions at 232 ms and 242 ms respectively (the time of the peak negativity in the error waveforms; units are μV).

The ERN for ignorance errors was significant ($F(1, 7) = 7.67$, $MSE = 9.55$, $p = 0.028$) while for violation errors it was not ($F(1, 7) = 1.13$, $p = 0.32$). The difference in the ERN between ignorance and violation errors approached significance ($F(1, 7) = 4.30$, $MSE = 4.98$, $p = 0.077$). This suggests the ERN is observed when participants have no knowledge of the association, but is not observed when they have some knowledge of the mapping but their expectations are violated.

The design of this experiment also allowed investigation into ERPs following value presentation on error trials. Because error feedback means that participants selected the wrong value, the actual value presented on error trials will be either better or worse than

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expected. Figure 2.11 shows the waveform following presentation of the expected outcome (i.e., correct feedback) in comparison with outcomes that are better or worse than expected (again only trials with the three highest confidence ratings are included, which indicates some degree of expectation on the part of the participant).

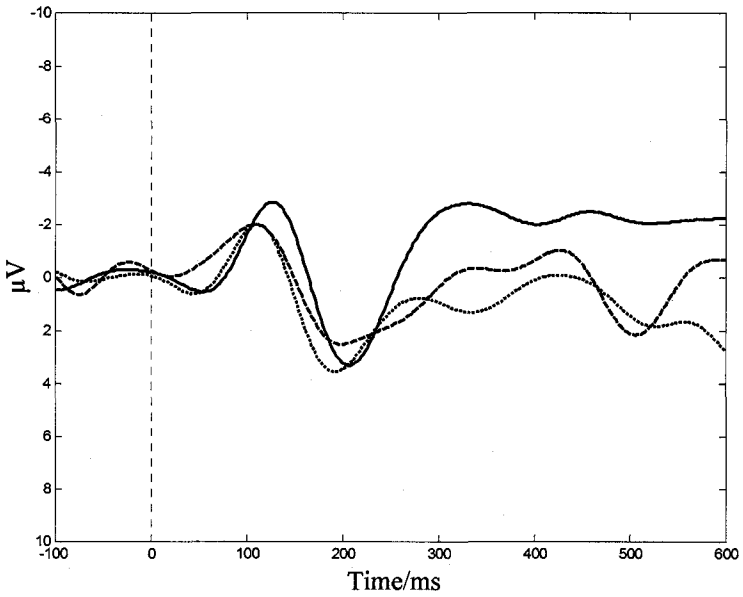


Figure 2.11 Experiment 1a waveforms following presentation of value stimuli. Waveforms for values that were expected (solid line), better than expected (dashed line) and worse than expected (dotted line) at electrode FCz.

No ERN-like activity was observed for outcomes that were not expected, nor did there appear to be any difference between the waveforms following better or worse than expected outcomes. There was, however, a clear difference at about 300 ms between the waveform following the expected outcome and the waveform following either unexpected outcome. This is probably due to an overlapping P300 component usually having a more parietal distribution and frequently observed following unexpected events.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Experiment 1b

Design and Rationale

Experiment 1b was essentially the same as Experiment 1a, with two minor modifications to the probability mappings. One change was that all the mappings were the same strength. This avoided the problem of confounding frequency with different parts and confidence groupings. The other change was that a condition was included that had an association with the -20 outcome value. Experiment 1a did not find an electrophysiological effect of outcome value. One possibility for this might have been that a large enough loss was not included; so it was hoped that including an association with a larger loss might evoked more value-related ERP activity.

Methods

Participants

There were 12 participants (3 males and 9 females) aged between 19 and 21. All were right-handed, had normal or corrected-to-normal vision and were students at the University of Michigan. They were again paid $8 per hour plus bonus money they accumulated over the course of the experiment (described below).

Procedure

The procedure in Experiment 1b was the same as Experiment 1a. The design parameters were identical but the associative relationships to be learnt were different. In Experiment 1b there were 4 probability mappings rather than 6 and each outcome value was associated with a stimulus 70% of the time. Again the remaining association probability

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was divided equally between the other three outcomes. This is shown (again with some example Attneave shapes) in Table 2.2.

Experiment 1b	Outcome Value				Sample Attneave Shape
	-20	**-2**	**+2**	**+20**	
Association Strength (%)	10	10	10	**70**	
	10	10	**70**	10	
	10	**70**	10	10	
	70	10	10	10	

Table 2.2 Association strengths for each outcome value in Experiment 1b and example Attneave shapes.

Again the participants completed 8 blocks of 90 trials in which 3 mappings were learnt. Therefore, in Experiment 1b each participant completed 6 blocks of each probability mapping.

Electrophysiological Methods

The electrophysiological methods in Experiment 1b were identical to those in Experiment 1a.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Results

Behavioral Results

In general, the behavioral results of Experiment 1b were similar to those of Experiment 1a, although accuracy and confidence were somewhat lower. This is illustrated in Figure 2.12, which shows increasing accuracy over the three parts of the block, although of a lower magnitude than Experiment 1a. A one-way repeated measures ANOVA showed there was a significant difference between the parts ($F(2,22) = 10.11$, $p = 0.0023$, $MSE = 0.016$). A subsequent Tukey HSD test revealed that accuracy in part 1 was significantly lower than that in part 2 ($p = 0.021$) and part 3 ($p = 0.0007$); however, the difference in accuracy between parts 2 and 3 was not significantly different ($p = 0.30$).

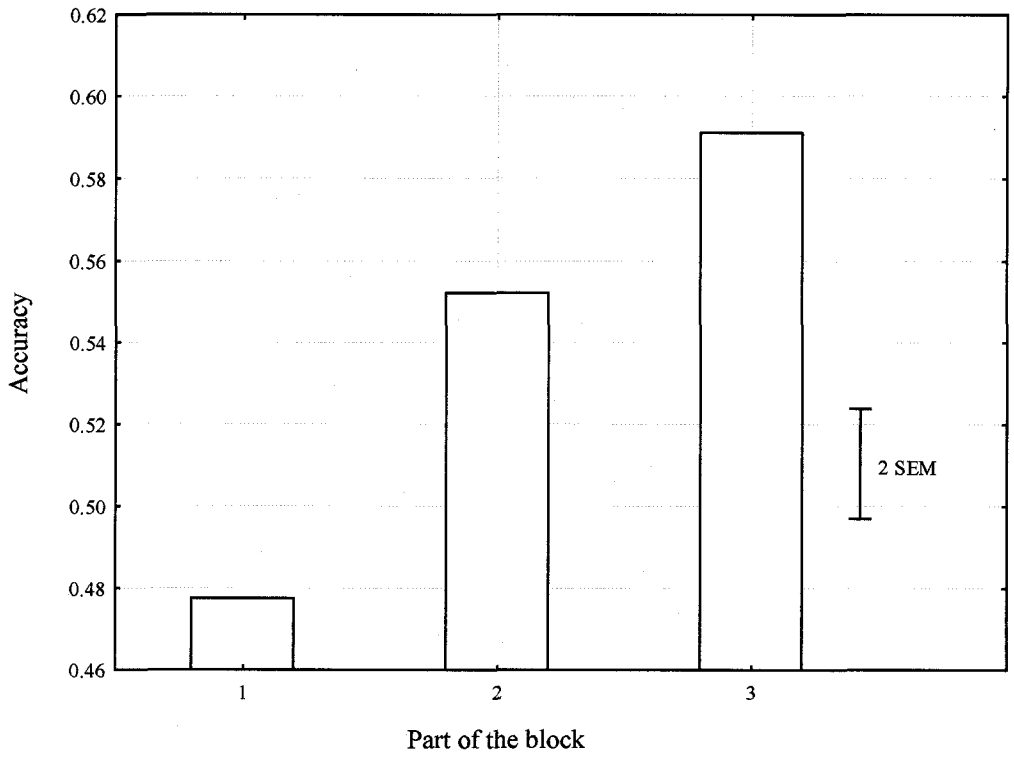


Figure 2.12 Experiment 1b change in accuracy over the block.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The results also showed (Figure 2.13) that there was a significant difference in accuracy between the four associated values ($F(3, 33) = 4.64$, $MSE = 0.012$, $p = 0.012$). This is surprising because all the outcome values had the same association with the stimuli (70%). Tukey HSD tests revealed that accuracy was significantly lower for the -20 value than the +2 value ($p = 0.011$) and the +20 value ($p = 0.024$).

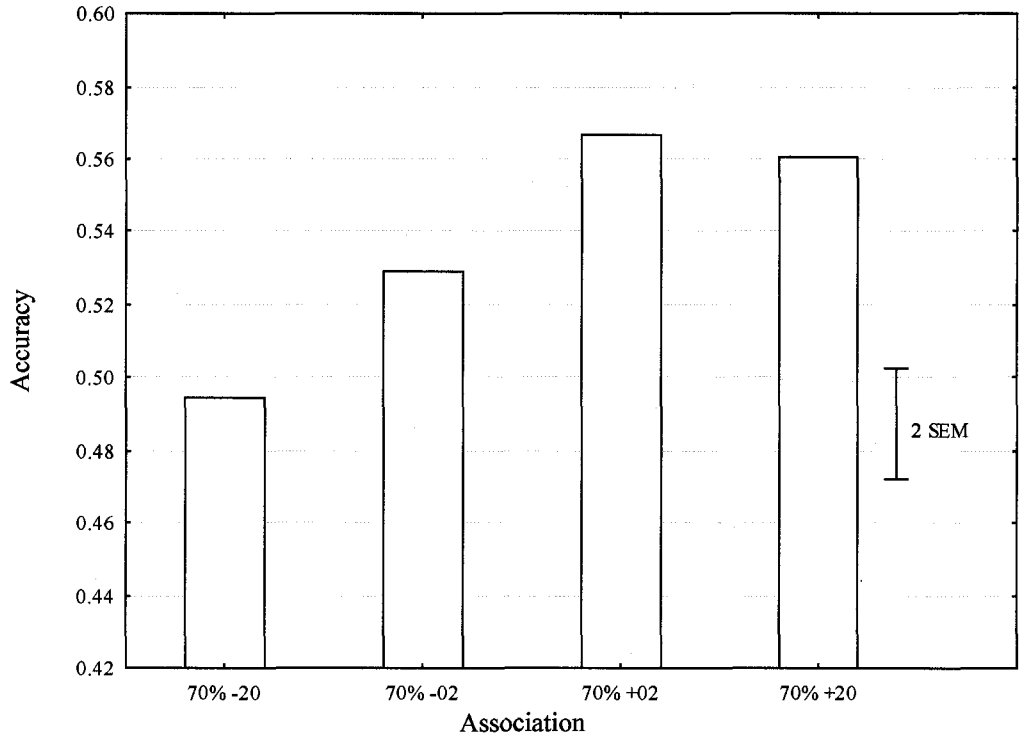


Figure 2.13 Experiment 1b accuracy for the different associations.

Also, as in Experiment 1a, there was a significant difference in accuracy between the different confidence ratings ($F(3,33) = 9.26$, $MSE = 0.049$, $p = 0.004$). Figure 2.14 shows accuracy increased with confidence until the highest confidence rating. A Tukey HSD test revealed accuracy at the low confidence rating was significantly lower than accuracy at the medium-low rating ($p = 0.0028$) and the medium-high rating ($p = 0.00025$), but the difference between accuracy at low and high confidence ratings was only marginally significant ($p = 0.056$). The low accuracy and greater variability for high

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

confidence ratings is probably due to the general low confidence in this experiment and the resulting low number of high confidence ratings made (280 compared with over 2000 for each of the other ratings). In fact, a 70% association strength might only be expected to result in few (if any) high confidence judgments.

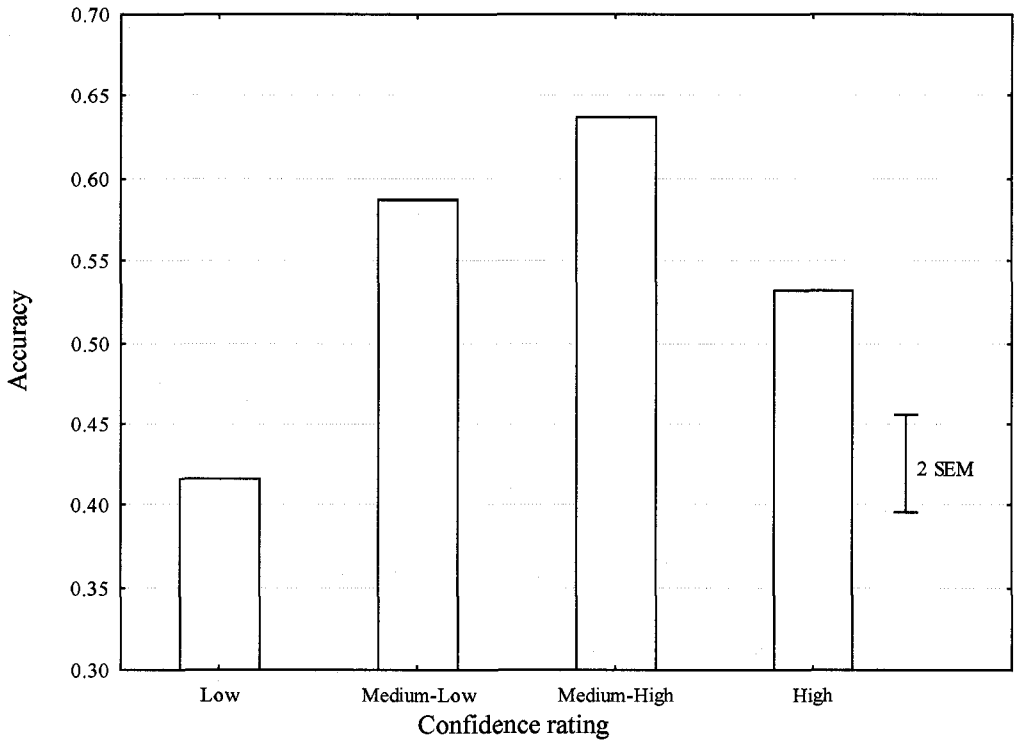


Figure 2.14 Experiment 1b accuracy for the different confidence ratings.

As in Experiment 1a, confidence ratings showed a similar pattern to accuracy. Confidence increased over the three parts of the block ($F(2, 22) = 16.18$, $MSE = 0.054$, $p < 0.0001$). However, confidence ratings in general were much lower than in Experiment 1a (2.7 in Experiment 1a versus 1.9 in Experiment 1b on a 1-4 scale). There was no significant difference in the confidence ratings of the different outcome values (all associated with shapes at 70%; $F(3, 33) = 2.36$, $MSE = 0.022$, $p = 0.11$).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Electrophysiological Results

On the whole, the electrophysiological results in the Experiment 1b were similar to those in Experiment 1a, although the differences between conditions were not as great. Also, while the time-course of the waveforms was very similar, their general morphology was somewhat different. The ERN observed in the Experiment 1b seemed to be extended over time compared with those in the Experiment 1a (see the end of Chapter 3 for a brief discussion). This can be seen in Figure 2.15, which shows the overall error and correct waveforms ($F(1, 11) = 14.72$, $MSE = 2.74$, $p = 0.0028$). The error waveform (in the 200-300 ms window, where the ERN is usually observed) peaked at 234 ms (very similar to the 236 ms peak observed in Experiment 1a). Again, similar to Experiment 1a, the topographic map of the ERN shows the negativity focused at medial-frontal electrode sites.

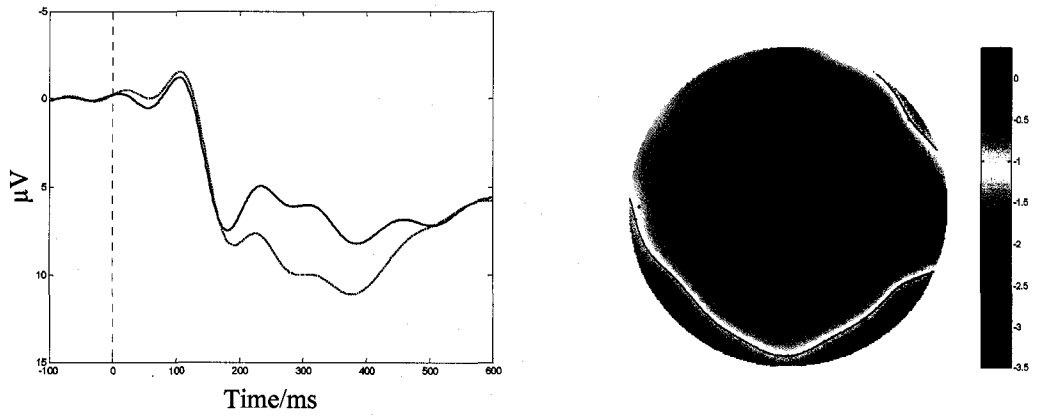


Figure 2.15 Experiment 1b ERN averaged over all conditions. Again, the error waveforms are solid and the correct waveforms are dotted. The topographic map is plotted at 234 ms (units are μV).

The change in ERN over the course of the block is shown in Figure 2.16; again, its amplitude appears to decrease over time. This was reflected by a significant interaction in a 3 (part) x 2 (accuracy) ANOVA ($F(2, 22) = 3.75$, $MSE = 1.15$, $p = 0.041$). While

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

there was a significant difference between the error and correct waveforms for each part (early: $F(1, 11) = 22.11$, $MSE = 3.50$, $p = 0.0006$; middle: $F = (1, 11) = 8.54$, $MSE = 2.92$, $p = 0.014$; late: $F = (1, 11) = 8.85$, $MSE = 3.38$, $p = 0.013$), the ERN (the error-correct difference) was significantly larger in the early part than the middle part ($F = (1, 11) = 6.69$, $MSE = 1.08$, $p = 0.025$) and marginally significantly larger than the late part ($F(1, 11) = 4.32$, $MSE = 1.28$, $p = 0.062$). The difference between the middle and late parts was not significant ($p = 0.75$).

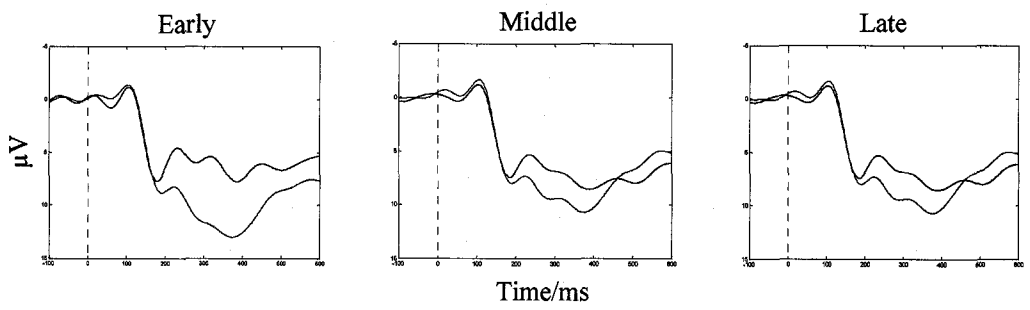


Figure 2.16 Experiment 1b change in ERN over the block.

The error and correct waveforms for the different confidence judgments showed results similar to those found in Experiment 1a (Figure 2.17).

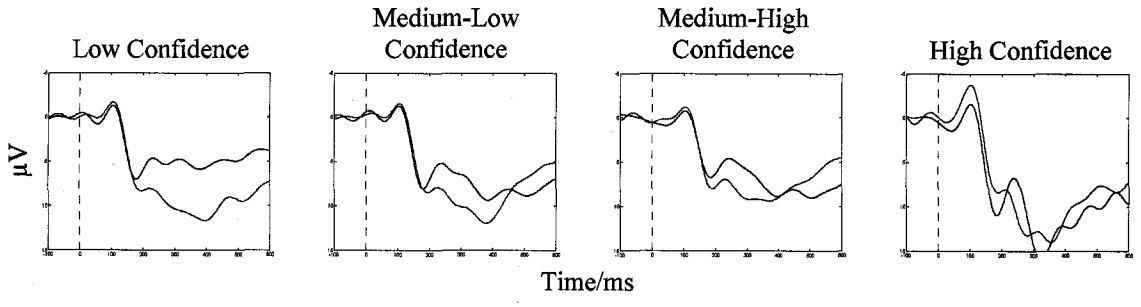


Figure 2.17 Experiment 1b ERN for the four confidence ratings.

There was a significant ERN in all but the highest confidence rating (low: $F(1,11) = 18.28$, $MSE = 3.68$, $p = 0.0013$; medium-low: $F(1,11) = 8.94$, $MSE = 3.89$, $p = 0.012$; medium-high: $F(1,11) = 9.21$, $MSE = 3.48$, $p = 0.0011$). However, no ERN for any of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the confidence ratings was significantly different from any other ($p > 0.20$). As in Experiment 1a these confidence ratings can be collapsed into two rating groups, low and high. However, combining the groups in the same manner as Experiment 1a might not be appropriate given the overall lower confidence in Experiment 1b. In Experiment 1a the two lowest confidence ratings were collapsed into the low group and the two highest were collapsed into the high group. Based on the raw number of confidence ratings (low: 1195, medium-low: 1326, medium-high: 840, high: 1941) this seems appropriate because there are a roughly equivalent number of trials in the collapsed low and high groups. The raw number of responses in Experiment 1b (low: 3208, medium-low: 2813, medium-high: 2001, high: 280) suggests that a similar division might not be appropriate because the two groups would have widely differing numbers of trials. Therefore, the low group remains the same and the high group combines the medium-low, medium-high and high groups (this captures the overall lower confidence participants had in Experiment 1b compared with Experiment 1a). These combined waveforms are shown in Figure 2.18. A 2 (confidence) x 2 (accuracy) ANOVA showed there was a significant difference between the error and correct waveforms for both low ($F(1,11) = 18.28$, $MSE = 3.68$, $p = 0.0013$) and high confidences ($F(1, 11) = 8.91$, $MSE = 2.92$, $p = 0.012$). However, the interaction was also significant, suggesting the ERN was greater for the low than the high confidence ratings ($F(1, 11) = 4.88$, $MSE = 0.99$, $p = 0.049$). The topographic maps again show a greater negativity at medial-frontal electrode sites for low confidence errors than high confidence ones.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

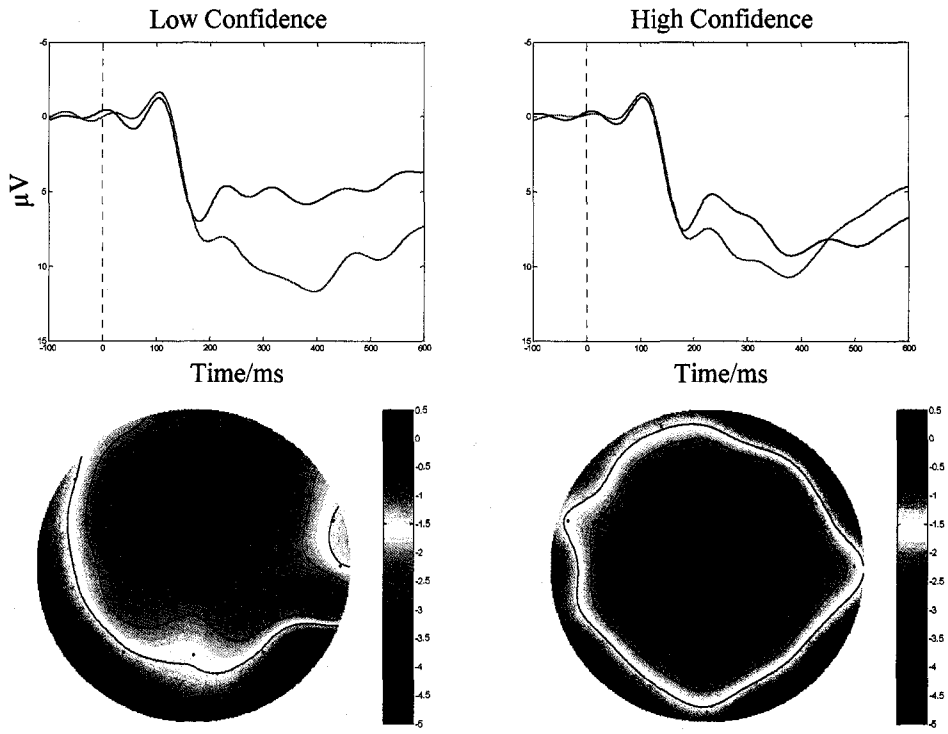


Figure 2.18 Experiment 1b ERN for grouped low and high confidence ratings. The lower panels show the topographic maps (units are μV)

Again there is a confound between part and confidence because most low confidence ratings will be made early in the block, so it instructive to look at the waveforms for error and correct trials for different confidence ratings at the different parts in the block (Figure 2.19, the low and high confidence groupings have been combined as described above). These results were somewhat different from Experiment 1a. There was a significant difference between the error and correct waveforms in the low confidence early part ($F(1, 11) = 19.52$, $MSE = 5.22$, $p = 0.001$) and in the high confidence middle ($F(1, 11) = 7.18$, $MSE = 4.82$, $p = 0.021$) and late parts ($F(1, 11) = 6.34$, $MSE = 4.41$, $p = 0.029$). There was no significant difference between the ERNs for any confidence rating in any of the parts.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

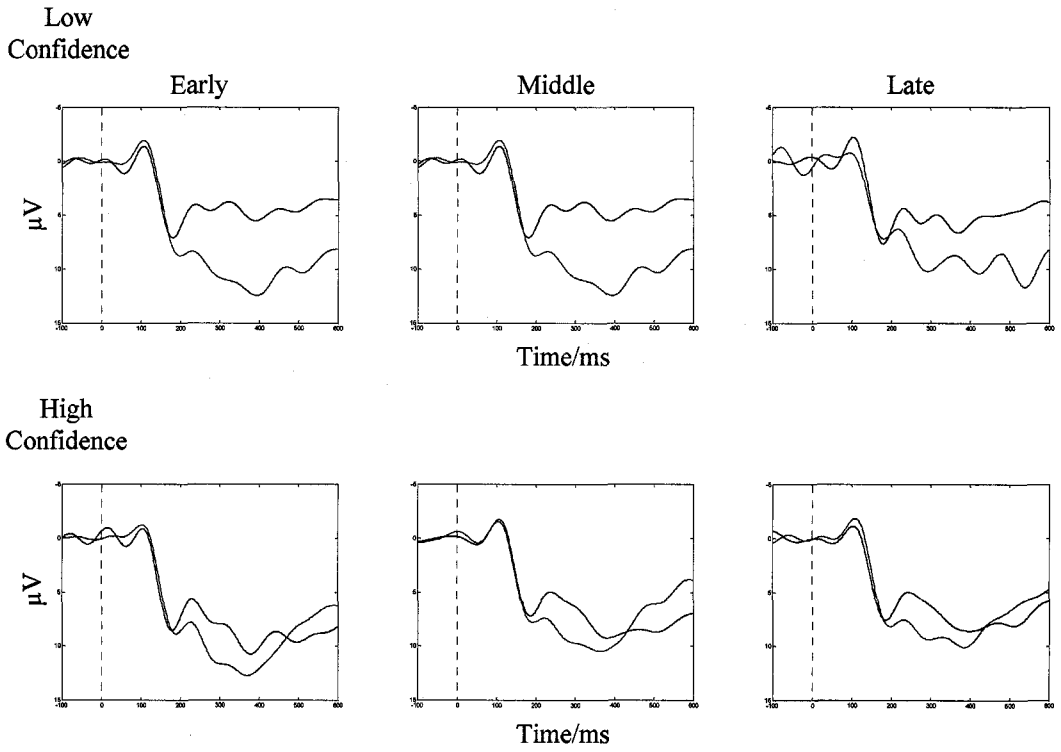


Figure 2.19 Experiment 1b change in ERN over the block for the grouped low and high confidence ratings.

Categorizing errors into ignorance and violation errors (using the same criteria as in Experiment 1a), again showed a similar pattern to Experiment 1a, although the difference between conditions was not as pronounced (Figure 2.20 shows the waveform plots and the scalp distribution of the ERN). A 2 (error type) x 2 (accuracy) ANOVA showed the difference between error and correct waveforms was significant for both ignorance errors ($F(1, 11) = 18.28$, $MSE = 3.68$, $p = 0.0013$) and violation errors ($F(1, 11) = 9.24$, $MSE = 3.19$, $p = 0.011$). The interaction approached significance ($F(1, 11) = 3.94$, $MSE = 0.98$, $p = 0.073$), suggesting the ERN was marginally larger following ignorance errors than violation errors. The topographic map illustrates the larger negativity at medial-frontal electrode sites for the ignorance errors, although it appears to have a more right-sided distribution than a typical ERN.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

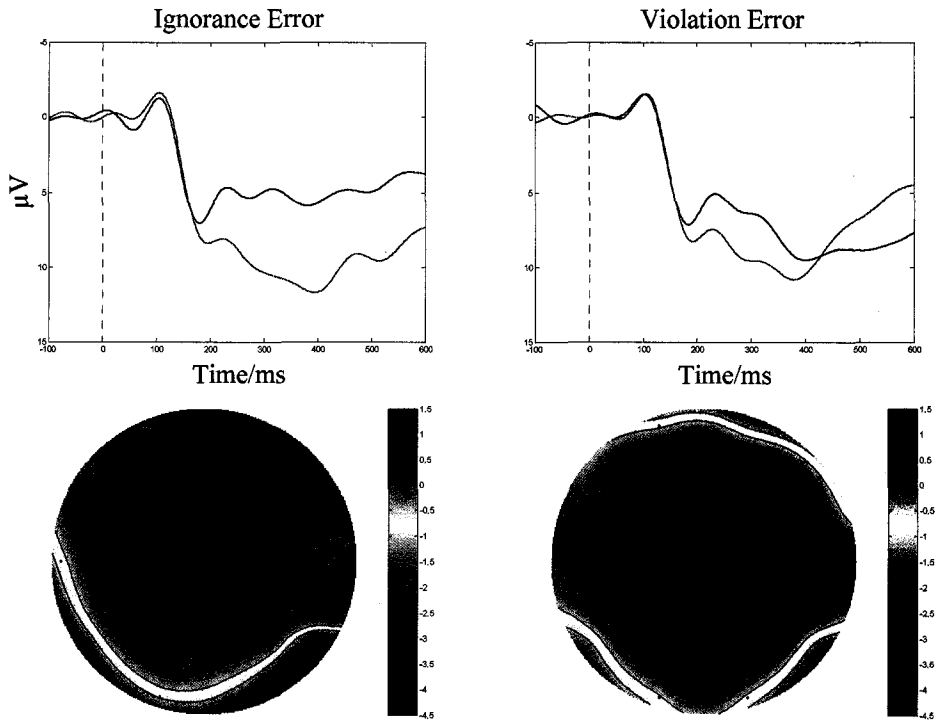


Figure 2.20 Experiment 1b ignorance and violation ERNs. The topographic maps are both at 234 ms (units are μV).

Turning now to the ERPs following the monetary outcomes, Figure 2.21 shows the waveforms when the outcomes were expected, or better or worse than expected. No ERN-like waveform was observed following presentation of value outcomes, nor did there appear to be a difference between waveforms following outcomes that were better or worse than expected. As in Experiment 1a, there appeared to be a difference at about 300 ms that can probably be attributed to the P300.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

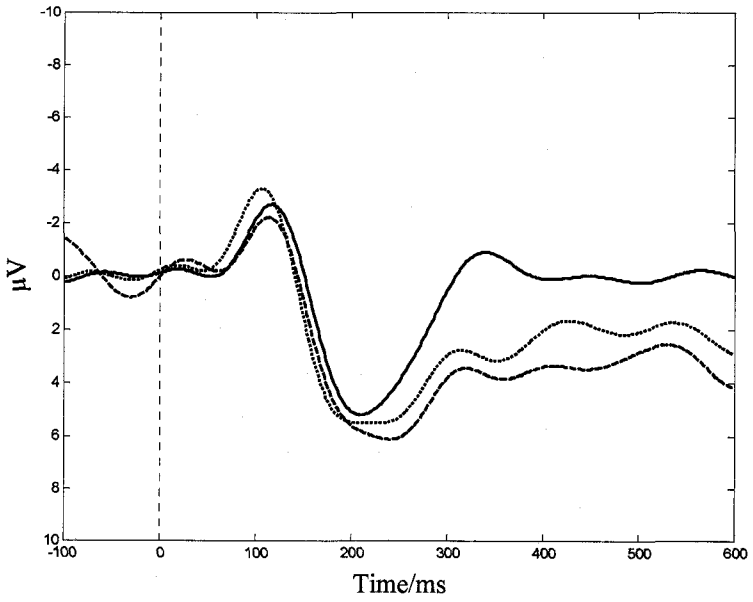


Figure 2.21 Experiment 1a waveforms following value presentation. Waveforms for values that were expected (solid line), better than expected (dashed line) and worse than expected (dotted line) at electrode FCz.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Experiment 1 Discussion

The most striking results from Experiment 1a were that the ERN appeared to decrease over the course of learning and with increasing confidence. While it is not entirely clear whether the learning itself or concomitant increase in confidence was more closely associated with this decrease, the results suggest the decrease in ERN was more strongly related to the confidence participants had in their choices rather than actually how long they had been learning the associations. Furthermore, a related analysis that separated ignorance errors (when participants had not learnt the association) from violation errors (when they had learnt the association, i.e., violations of expectancy) showed the ERN was larger following ignorance than violation errors. Similar results were observed in Experiment 1b, although the differences between the conditions were not as great. One potential explanation for this could be that because confidence was lower in Experiment 1b than 1a (possibly resulting from the different probability mappings used) and because the ERN decreases with increasing confidence, it would not have decreased in Experiment 1b as much as it did in 1a.

These results do not appear to be consistent with the prediction made by the RL-ERN. This model predicts the ERN would be largest when strongly held expectations are violated, that is, late in learning or following violation errors. However, as the results of Horst et al. (1980), Kopp and Wolff (2000) and Butterfield and Mangels (2003) discussed earlier suggest, findings such as those in the present study are not unprecedented. All of these studies found a much stronger relationship between participants' violation of expectancy and the P3/P300 than the ERN.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

There are, however, two other studies (very similar to each other) designed to test the RL-ERN theory (Holroyd and Coles, 2002; Nieuwenhuis et al., 2002). In Holroyd and Coles' (2002) two-choice reaction time experiment, stimuli of six different categories were associated with one of six probability associations (always correct, always incorrect, two 100% conditions and two 50% conditions). There was also a reaction time deadline that ensured impulsive errors in the 100% conditions even after the association had been learnt, thereby enabling a comparison between response and feedback ERNs. It was predicted that the first event to provide performance feedback would elicit a larger ERN. For example, after the 100% mapping was learnt, the response ERN would be expected to be largest because the system would be able to detect the error (a worse than expected event) immediately after it had been committed. However, for the random condition, the system would not be able to recognize immediately whether the response was correct, so it would have to wait for accuracy feedback (when there would be a larger ERN following error feedback). The experimental results lent some support to the model; the response ERN increased over learning in the 100% and the feedback ERN was larger than the response ERN in the 50% condition. However, some other predictions, such as a decrease in the feedback ERN in the 100% condition, were not supported.

The task employed by Nieuwenhuis et al. (2002) was very similar to that used by Holroyd and Coles (2002). However, instead of having 'always correct' and 'always incorrect' conditions, they included two 80% conditions where after learning feedback was unexpected on 20% of the trials. In accordance with the model's prediction the largest feedback ERN was observed following unexpected error feedback. This result is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in contrast to the findings of Experiments 1a and 1b, which suggests the ERN is not related to violations of expectancy.

The most tempting explanation of the results reported here is that the ERN is related to learning about the probability mapping from the error feedback. That is, when the mapping is not known, error feedback provides information about the mapping. However, after the mapping has been learnt, error feedback provides no more information about the mapping, so the amplitude of the ERN is reduced (even if outcomes on individual trials are worse than expected). Contrasting this explanation with the predictions of the RL-ERN model highlights different kinds of expectation a participant may have, namely, expectations about the outcome of a particular trial and expectations about the distributions of outcomes over a number of trials. For example, consider the condition in which there is a strong association between the stimulus and a gain of 20 cents, the expectation held could be that there is an 85% chance of winning 20 cents on each individual trial. However, the expectation could alternatively be that of definitely winning 20 cents on 85% of the trials.

A problem with the RL-ERN (and connectionist models in general) is that it is difficult to tell how information is represented across a population of units; the expectation could be represented in either (or, potentially, both) of these two fashions. A closer look at the RL-ERN model, however, may reveal some clues as to how it represents expectations. The RL-ERN incorporates a summation unit in the adaptive critic whose activity may be a good candidate for the representation of the probability of an outcome following a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

particular response on a particular trial. Furthermore, the learning dynamics of the two different representations would also differ. The effect of a single unexpected outcome on the individual trial representation would be much greater than on a distribution representation; therefore, the expectation derived from an individual trial representation would change more quickly. Results of sequential trial analysis in Holroyd and Coles' (2002) study suggest that the model updates its expectations immediately following an unexpected outcome. These clues suggest that the RL-ERN model represents expectation in terms of probability of a particular outcome on a particular trial.

While this interpretation involving information about the association is attractive, it is important not to dismiss others. A related explanation (that could be considered the affective component, or extension of, the previous one) may be that after participants have learnt the mapping they no longer feel as though they are making a mistake because they know they made the most appropriate response and, therefore, there is no motivation to change their behavior or update their expectations. Another possibility might be that the ERN reflects a more general arousal signal that motivates the participant to attend to the following outcome value stimulus because it is important for learning. In light of the literature reviewed earlier implicating the ACC and ERN in emotional and motivational functions these possibilities will be explored in more detail later.

Another notable finding was that no ERN-like component was observed following loss feedback. An explanation for this might be that participants paid little attention to the actual value of the outcomes because they were more concerned with learning the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

associations (as the task required them to do) and choosing the most appropriate responses.

While the results from Experiment 1 are both theoretically interesting and challenging there are still a number of unresolved issues. One issue concerns whether the changes in ERN amplitude vary predictably for different probability mappings. If the interpretation that the ERN is related to learning from the feedback stimulus is accepted, the ERN amplitude should decrease more slowly for weaker associations (that take longer to learn). A further issue concerns the need to dissociate expectancy of the mapping as a whole from expectancy of the outcome on an individual trial. Furthermore, additional data would enable stronger conclusions to be drawn regarding the relationship between confidence, learning and the ERN.

One final issue concerns the frequency of error feedback stimuli and the impact of the oddball N200, which frequently occurs with the P300. The N200 is characterized by a similar latency and scalp distribution to the feedback ERN and is typically observed following infrequent events. Therefore, if errors are infrequent events, it is important to know whether and to what extent the N200 influences the feedback ERN.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter III

Experiment 2

Design and rationale

Experiment 1 uncovered some interesting phenomena that can potentially help clarify the functional significance of the ERN. In particular, the decrease in ERN amplitude over the course of learning and the larger ERN following ignorance errors than violation errors suggest that the ERN might reflect a psychological process only present in the early stages of learning, which would have implications for accounts of ERN function. However, some experimental manipulations (such as the value of the outcome) had little or no effect. Furthermore, while the results across Experiments 1a and 1b were generally consistent, they did not correspond as closely as might have been expected. This could have been caused, at least in part, by the different probability mappings in the two experiments. Another potential consequence of using different probability mappings in Experiments 1a and 1b was that some of the results from Experiment 1a could not be replicated in Experiment 1b (because all the probability mappings were all the same strength). Moreover, for some of the more intriguing results, there were too few trials in the averaged waveforms (because the experiment was not designed to investigate these effects). Therefore, while similar in nature to Experiments 1a and 1b, Experiment 2

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

focused on the change in ERN over the learning of different probability mappings using a slightly modified experimental procedure.

One of the major changes implemented in Experiment 2 was that the stimuli were associated with responses rather than outcome values. The results from Experiments 1a and 1b show that there was no ERN-like ERP following the presentation of the outcome value. This suggests participants might have been more concerned about performing well in the learning task than the amount of money being gained or lost. Removal of value from the association to be learnt allowed it to be included as a reward or penalty for correct or incorrect responses. This afforded two benefits, one was that rewards and penalties were used to motivate appropriate task performance and the other was that a further probabilistic association could be investigated, that between accuracy feedback and reward or penalty value (in addition to the association between the stimulus and response explicitly learnt). One concern, given the results of Experiment 1, was that participants would pay more attention to the learning task than to the reward and penalty values; however, separating value feedback from the learning task in Experiment 2 potentially allows participants to pay more attention to the actual value.

Another difference in Experiment 2 was that only accuracy feedback was given, rather than presenting the correct response. This had two advantages: first, it avoided confounding different types of information in the feedback stimulus (for example, presenting the correct answer confounds accuracy information with information about the correct S-R mapping) and second, because only correct responses provide information

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

about the S-R mapping, the associations should take longer to learn, which in turn, should result in a broader distribution of confidence judgments.

A change in hardware also enabled a previously unforeseen and potentially beneficial methodological modification. Experiment 2 used a five-button response pad rather than the four-button pad with used in Experiment 1. This enabled a different mapping of confidence judgments to response. Rather than representing merely different levels of confidence that the previous choice was correct (as was the case in Experiment 1), participants were able to report the confidence with which they thought they had made either a correct or an incorrect response. As shown in Figure 3.1, the central button represented uncertainty in the previous choice, the two buttons to the left represented increasing confidence that they had made the incorrect choice and the two buttons to the right represented increasing confidence that they had made the correct choice. Therefore, both expected accuracy and confidence were measured.

Other methodological changes were that the block was not divided into three arbitrary parts. Instead, behavioral measures of accuracy were used to assess learning progress. Also, participants learnt the S-R mappings individually (rather than in groups of three as in Experiment 1). This change was made to simplify the task and decrease the influence of other psychological processes.

One of the goals of Experiment 2 was to separate the effects of learning and confidence. Because they are associated and highly correlated, finding a larger ERN early in learning

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

might be due to either not having learnt the association or low confidence. In order to separate these two factors, the crucial trials are the low confidence correct and high confidence error trials. Increasing the number of random and weak associations enabled more low confidence trials to be collected and continuing to present stimuli in the strong association condition after the mapping had been learnt allowed more high confidence error trials to be collected. low confidence error and high confidence correct trials are frequent because participants tend to make mistakes when they have not learnt the mapping when their confidence is low and tend to be both accurate and confident when they have learnt the association.

Given the results of Experiments 1a and 1b some predictions about the results of Experiment 2 can be entertained. The amplitude of the ERN decreased over the course of the block in Experiment 1; therefore, a similar decrease in amplitude might be expected in Experiment 2. Experiment 1a also used varying association strengths; however, the results did not show clearly whether the ERN decreased for all the association strengths. If the ERN is, indeed, related to learning it might be expected that it will decrease more slowly for those associations that are harder to learn.

There was a slight suggestion in Experiment 1 that the ERN might be more sensitive to confidence than how well the association has actually been learnt. While the evidence is not strong, if this is indeed true, Experiment 2 should show a larger ERN for low confidence trials and a smaller ERN for high confidence trials regardless how long has been spent learning the association (i.e., the position in the block). If, however, the ERN

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is sensitive to time learning the ERN should decrease over the course of the block regardless of confidence.

One issue that frequently complicates the analysis of the feedback ERN is overlap by the oddball N200/P300 component (Holroyd, 2004). This component is usually observed following infrequent events, which, in many cases, is the case for errors and error feedback. The N200 component shares similar characteristics with the feedback ERN – it has a medial-frontal scalp distribution, the same polarity and a similar time-course to the ERN – and there is some speculation that the N200 and the feedback ERN are the same component (e.g., Holroyd, 2004). The P300 is often tightly coupled with the N200 but has a more posterior scalp distribution and occurs slightly later. However, it still frequently overlaps the ERN. The probability mappings in Experiment 2 were selected to allow comparison of the ERN following error and correct feedback that was equally likely. If the ERN is insensitive to stimulus frequency an equal ERN might be expected to all error stimuli (at the same position in the block or followed by the same confidence rating). If, however, it is solely sensitive to stimulus frequency, a larger ERN might be expected following any infrequent stimulus, including infrequent correct feedback.

The final prediction concerns the association between the feedback stimulus and the value of the reward and penalty. By varying reward and penalty value, this experiment was also able to investigate the effects of violations of expectancy of reward value following accuracy feedback (as opposed to that of accuracy feedback following the response). The distribution of value feedback means that high and low rewards and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

penalties will be relatively unexpected, which enables investigation into the effects of not only simply gaining or losing money but also receiving rewards and penalties that are better or worse than expected. If the ERN functions as the RL-ERN suggests, it should only be sensitive to the degree to which the value is worse than expected, regardless of the actual monetary outcome. If, however, the ERN is sensitive to 'negative utility' events (e.g., losing money) it should be larger following losses regardless of participants' expectations.

Methods

Participants

There were 13 participants (1 male and 12 females) aged between 18 and 22. All were right-handed, had normal or corrected-to-normal vision and were students at the University of Michigan. They were paid $8 per hour plus bonus money accumulated over the course of the experiment (described below).

Procedure

The basic design of Experiment 2 is similar to that of Experiments 1a and 1b in that participants had to learn a probabilistic S-R mapping (i.e., which response to make when a stimulus was presented). Stimuli and responses were associated with three different strengths (as shown in Table 3.1). First, the participant was presented with an imperative stimulus (the fixation cross changed from white to red), following which the participant pressed one of four buttons. Then, the participant indicated their confidence in the accuracy of their previous response by pressing one of five buttons. As described before,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the center button represented uncertainty whether the previous response was correct (unsure), the two buttons to the left represented increasing confidence that the response was incorrect (possibly incorrect, probably incorrect) and the two buttons to the right represented increasing confidence that the response was correct (possibly correct, probably correct). Next, accuracy feedback was presented, informing the participant whether or not they did, in fact, make the correct response. This feedback was either the @ or the # symbol, the meaning of which (error or correct) was counterbalanced across subjects. Finally, feedback was presented indicating the value of monetary reward or penalty (5, 25 or 50 cents), depending on whether or not they made the correct response. The value 25 was presented 50% of the time and 5 and 50 were each presented 25% of the time. It was hoped that the relative frequency of the values would allow further investigation into ERP components related to the presentation of better or worse than expected outcomes. A schematic illustration of Experiment 2 is presented in Figure 3.1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

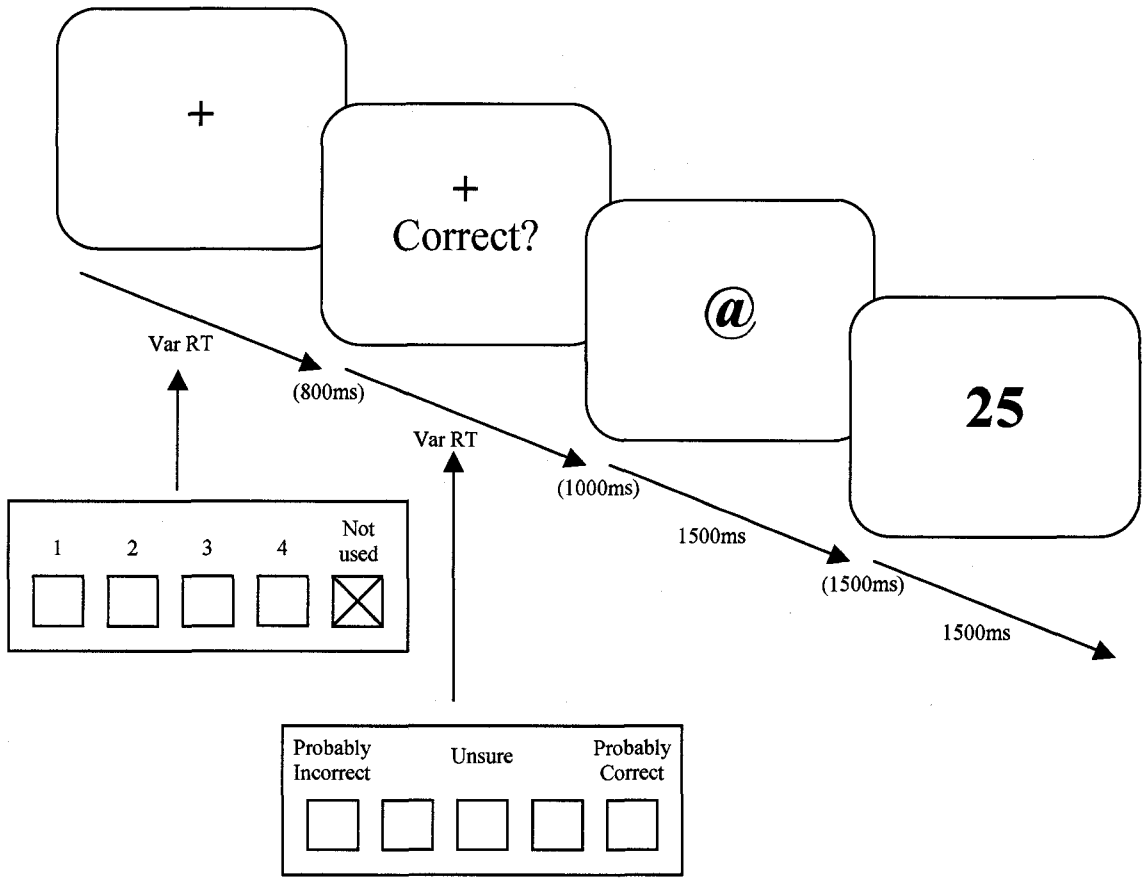


Figure 3.1 A schematic illustration of Experiment 2.

The probability mappings are shown in Table 3.1. These were chosen for two reasons. First, the strong association condition would still elicit errors on 25% of the trials after the association has been learnt, which would enable collection of a reasonable number trials in which unexpected error feedback was given. Second, after the strong association has been learnt, error feedback will occur with the same probability as correct feedback in the random condition (i.e., on 25% of the trials). The opposite is also true, that correct feedback in the strong condition (when learnt) will occur with the same frequency as error feedback in the random condition (i.e., on 75% of the trials). Furthermore, when the weak probability mapping has been learnt, both error and correct feedback should occur with approximately the same frequency (50%). This should enable the contribution

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the oddball N200 and P300 to be assessed because error and correct feedback occurring with equal frequency can be compared.

Experiment 2		Response			
		1	**2**	**3**	**4**
Association Strength (%)	Strong	**75**	8.33	8.33	8.33
	Weak	**50**	16.66	16.66	16.66
	Random	25	25	25	25

Table 3.1 The association strengths in Experiment 2.

The order of presentation of the association strengths was randomized, as was the button assigned as the correct response. Participants completed 8 blocks of 40 trials for each of the three association strengths; therefore, they completed 24 blocks in total, which took approximately three hours.

Electrophysiological Methods

The electrophysiological methods used in Experiment 2 were very similar to those used in Experiment 1. All recording and analysis parameters were the same with two exceptions. Thirty-one (rather than 37) scalp electrodes were used, positioned at sites, Fp1, Fp2, AFz, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, O1, Oz and O2 (Jasper, 1958). Vertical eye movements were measured using a pair of electrodes above and below the left eye only.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Results

Behavioral Results

The basic behavioral results of Experiment 2 replicate those of Experiment 1. Accuracy increased over the course of the block, was higher for the stronger association strengths and increased with confidence. Figure 3.2 shows change in accuracy for the three association strengths over all 40 trials.

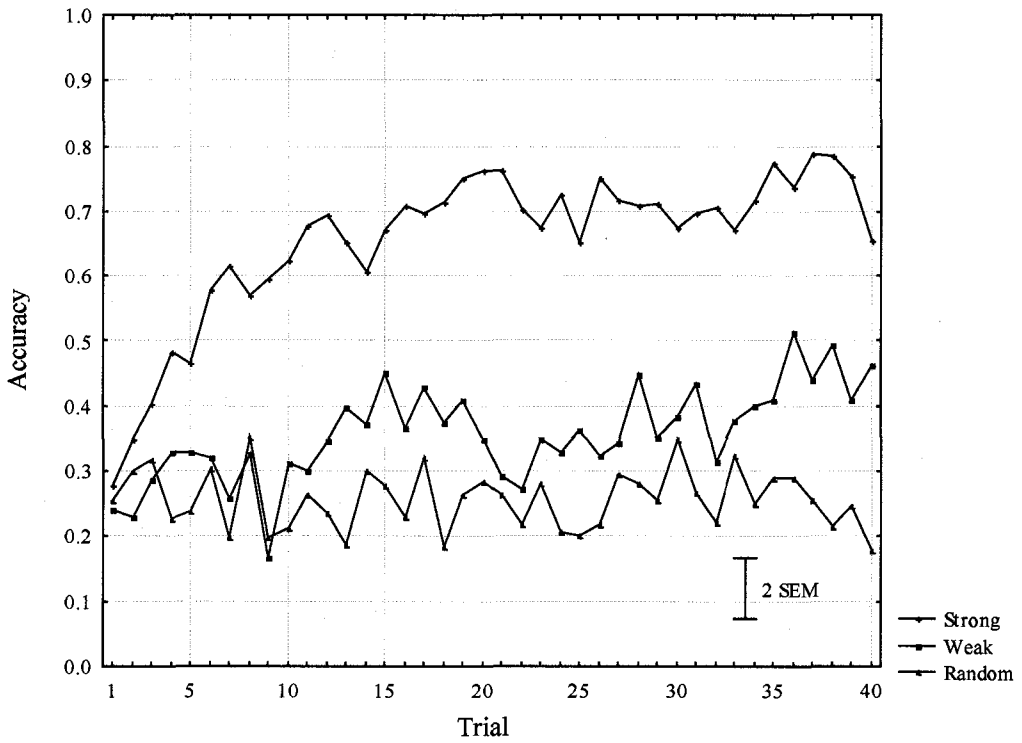


Figure 3.2 Experiment 2 learning curves for the three association strengths. As in Experiment 1, error bars show two standard errors of the mean.

Inspection of Figure 3.2 suggests learning the strong association is virtually completed within the first quarter of the block. Figure 3.3 shows the averages for each quarter and association strength. A 3 (association strength) x 4 (quarter) ANOVA revealed a significant main effect of association strength ($F(2,24) = 321.096$, $MSE = 0.0309$, $p < 0.001$). As in Experiment 1, all statistical tests were carried out on arc sine corrected

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

proportion data. A post-hoc Tukey HSD test indicated that mean accuracy in last three quarters of the strong association condition did not differ significantly ($p = 0.94$) but the accuracy in the first quarter was significantly lower ($p < 0.001$).

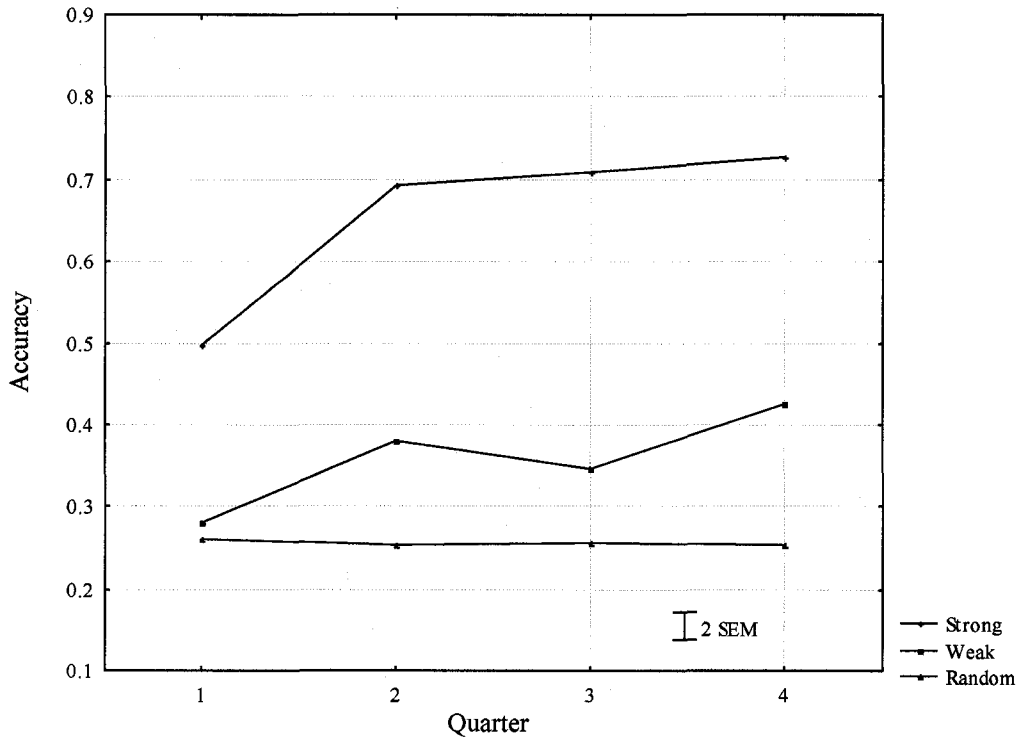


Figure 3.3 Experiment 2 learning over the course of the block for the three association strengths. Trials have been averaged into quarters.

Also, as in Experiment 1, a one-way ANOVA showed an effect of confidence ($F(4, 48) = 29.756$, $MSE = 0.0485$, $p < 0.0001$; Figure 3.4). However, the different confidence ratings in Experiment 2 reveal some interesting behavioral findings. While accuracy increased as the participant indicated an increased confidence in having made a correct response, accuracy did not decrease when participants indicated they had thought they had made an incorrect response. A post-hoc Tukey HSD test showed that accuracy on trials where the participant thought they had made an incorrect response did not differ from that on trials in which an unsure response was given ($p > 0.95$).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

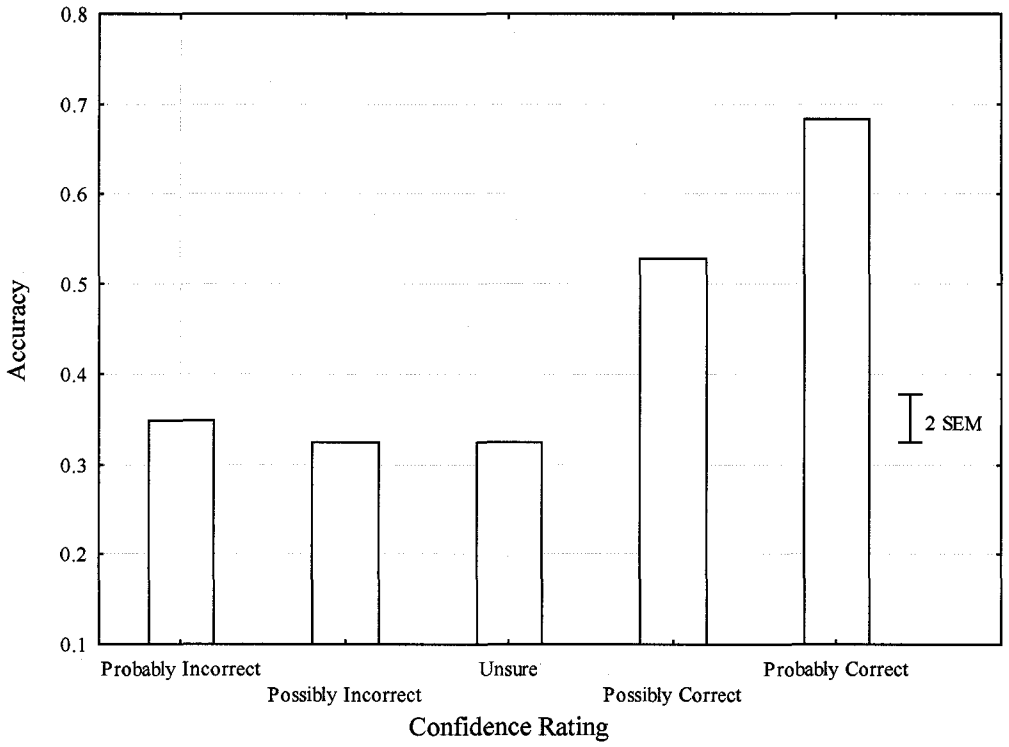


Figure 3.4 Experiment 2 accuracy for the different confidence ratings.

Participants' ratings of their confidence in their accuracy allowed investigation into how confidence changed over the block (Figure 3.5). The results suggest that subjects quickly indicated an increased confidence that their responses were correct on strong association blocks but appeared to be much slower to indicate increased confidence that their responses were incorrect during random blocks. As might be expected, confidence ratings during weak association blocks did not change because even after the mapping had been learnt, the probability of both correct and error feedback was 50%.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

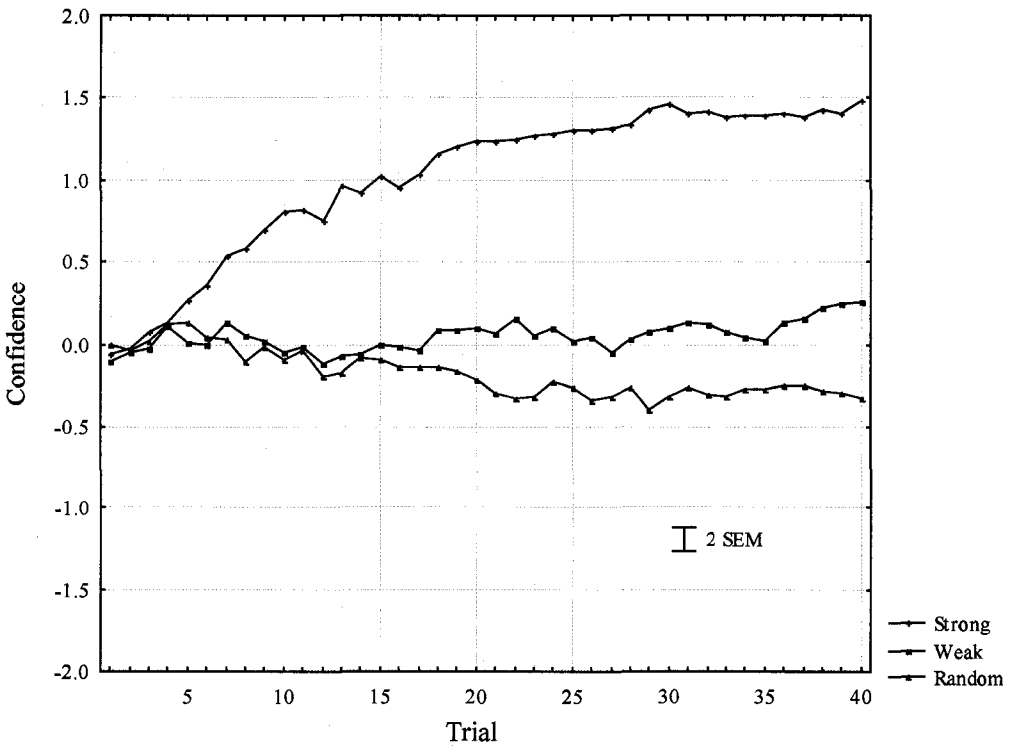


Figure 3.5 Experiment 2 confidence for responses for the different association strengths.

Also, as in Experiment 1, confidence ratings increased for stronger association conditions ($F(2, 24) = 38.264$, $MSE = 5.254$, $p < 0.0001$).

Electrophysiological Results

The results observed in Experiment 2 showed both similarities and differences with those found in Experiment 1. An ERN was observed following error feedback with a similar time course to that in Experiment 1 with the peak negativity of the error waveform at 240 ms; however, it had a more frontal distribution, being maximal at Fz rather than FCz (Figure 3.6). Therefore, unless otherwise stated, all analyses are carried out at electrode Fz. As in Experiment 1, all ANOVA probability values are Greenhouse-Geisser corrected where appropriate and waveforms are shown filtered with a Chebyshev type II low-pass filter with a half-amplitude cut off at 12 Hz.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

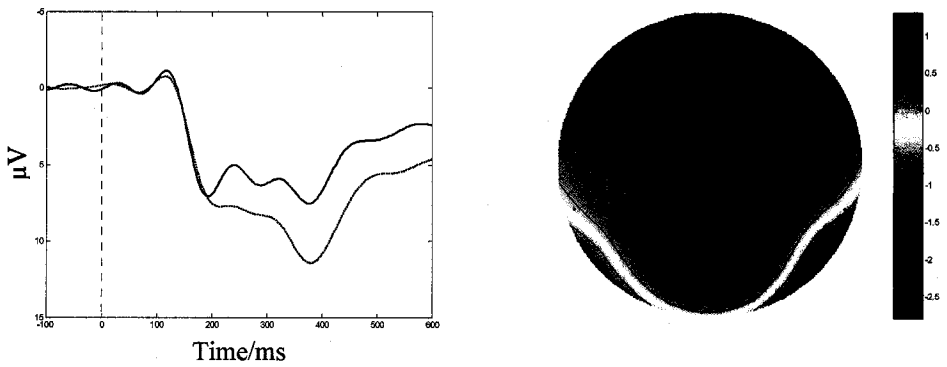


Figure 3.6 Experiment 2 ERN averaged over all conditions. As with the waveform plots for Experiment 1, error trials are plotted in solid lines and correct trials are plotted in dotted lines. The right panel shows the topographic map (units are μV).

The major finding from Experiment 1 was that the amplitude of the ERN decreased over the course of the block. In Experiment 1, the block was divided into thirds and the waveforms averaged by third. However, in Experiment 2 behavioral results, such as the learning curves for the different association strengths, were used to assess learning. They suggest that, as an approximation across all the association strengths, if the block is broken up into two parts, one comprising the first third and the other comprising the second two thirds, this reflects periods where, on the whole, learning is still in progress and learning has been completed. Figure 3.7 shows the correct and error waveforms for the first third of the block and the last two thirds averaged over all association strengths.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

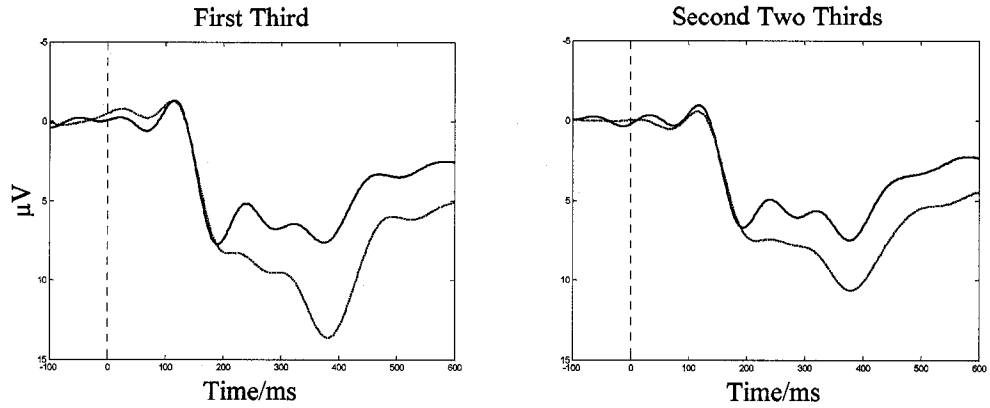


Figure 3.7 Experiment 2 change in ERN over the block.

An ERP is normally measured as the difference in amplitude between two waveforms (in this case the error and correct waveforms); however, this measure can be affected by differences in either waveform. This is acceptable if one of the waveforms is relatively stable across conditions, so the difference in the ERP across conditions can be attributed to the one varying waveform. However, considerable variation was observed in the correct waveforms in this experiment (e.g., Figure 3.8), so attributing differences in the ERN across conditions simply to variation in the error waveform is not possible. Experiment 2, therefore, employed a peak-to-peak measure of ERP amplitude in only the error waveform. While this technique does enjoy some benefits over a measure of the difference between two waveforms, it also has drawbacks of its own; for example, other components may influence the peak-to-peak measure because they are not subtracted out (as is the case when analyzing the difference between waveforms).

For the peak-to-peak measurement, the ERN (the negativity in the error waveform) was calculated as the difference between the negative peak in a 200-320 ms range and the positive peak preceding it in the 150-250 ms interval. This analysis for the error

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

waveforms shown in Figure 3.7 revealed that the ERN was significantly greater in the first third than in the second two thirds ($F(1,12) = 5.862$, $MSE = 0.7373$, $p = 0.032$). All comparisons presented here between error waveforms use peak-to-peak analysis, while comparisons between error and correct waveforms use mean amplitude measures in a 200-300 ms window following feedback presentation.

The ERN changes over the block might, however, vary for different association strengths. Therefore, the change in ERN over the block was examined separately for the different association strengths. The block was divided into two parts based on behavioral data such that learning was still taking place in the first part and was ostensibly completed in the second. Figure 3.2, showing the learning curves for the different association strengths, suggests the strong probability mapping is virtually learnt in the first quarter of the block. Therefore, the first quarter, in which learning is still occurring, was compared with the last three quarters, in which the mapping has been learnt, in order to compare the ERN during and after learning. This comparison is shown in Figure 3.8; the ERN in the first quarter was significantly greater than that in the last three quarters ($F(1,12) = 8.16$, $MSE = 15.70$, $p = 0.0145$).

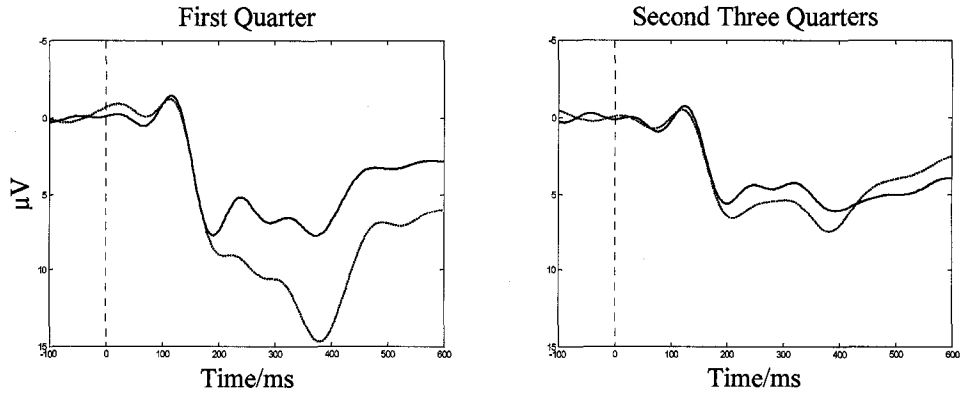


Figure 3.8 Experiment 2 change in ERN over the block for the strong association condition.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This division of the block, however, is not appropriate for the slower rate of learning of the other association strengths. Estimating the learning progress for the weak association strength is somewhat harder, because performance changes are not as pronounced as in the strong condition. Nevertheless, Figure 3.2 suggests that most learning took place by the second third of the block. Therefore, Figure 3.9 compares the error and correct waveforms for first third and the second two thirds of the block. While the negativity again appears to decrease, this effect is not significant ($F(1,12) = 0.736$, $MSE = 3.298$, $p = 0.408$).

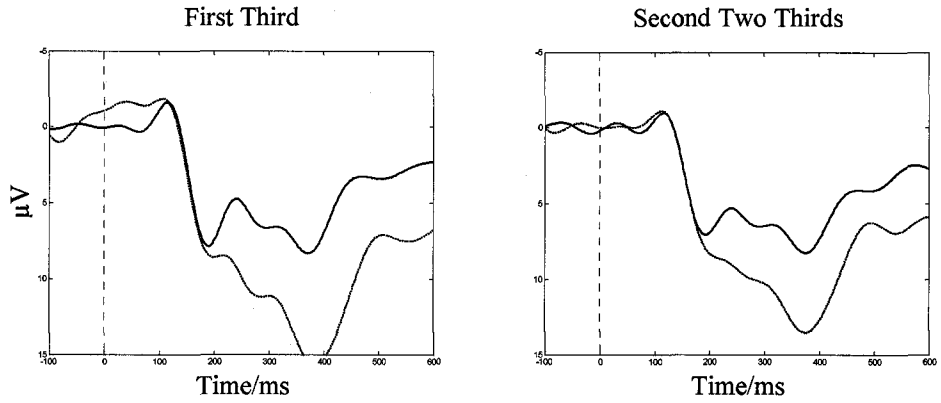


Figure 3.9 Experiment 2 change in ERN over the block for the weak association strength.

Accuracy during the random condition does not improve from chance; therefore, it cannot be used as a measure of learning. However, it is possible to estimate how well the association has been learnt by examining the change in confidence ratings. Figure 3.5, showing the difference in confidence ratings over the block, suggests that participants had learnt the random mapping approximately half way through the block. To reflect this Figure 3.10 shows the error and correct waveforms for the random association strength for the first and second halves of the block. Again the apparent decrease in negativity in the error waveform is not significant ($F(1,12) = 0.00047$, $MSE = 2.7413$, $p = 0.983$).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

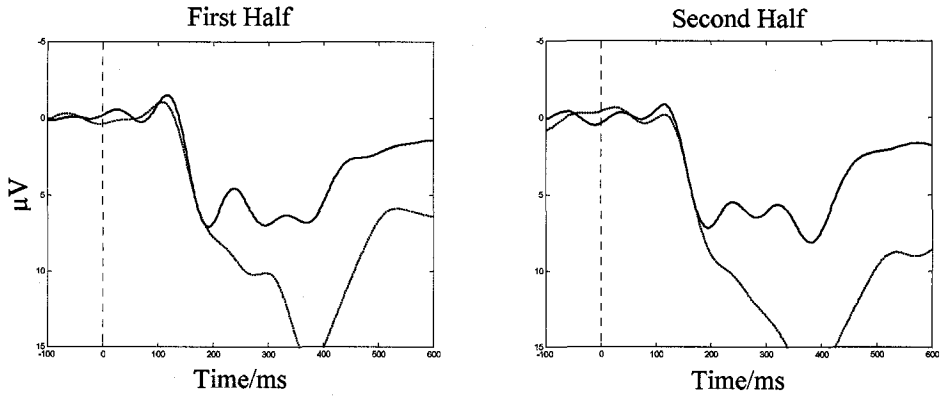


Figure 3.10 Experiment 2 change in ERN over the block for the random association strength

These comparisons between error waveforms averaged according to stage of learning seem to suggest that the decrease in negativity observed in the overall waveform is primarily driven by the decrease in negativity in the strong association condition.

Another feature of this experiment is that the waveforms can be grouped by confidence rating. Indeed, one of the objectives of Experiment 2 was to investigate whether the ERN was more closely related to learning the probability mapping or the confidence rating. Figure 3.11 shows the error and correct waveforms for different confidence ratings. These have been grouped into expected incorrect, unsure and expected correct. The expected incorrect and expected correct groups combine trials in which participants displayed some confidence in their response, either that it was incorrect or correct (in each case the averages comprised the 'possibly' and 'probably' ratings) or the unsure group simply contains the 'unsure' rating.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

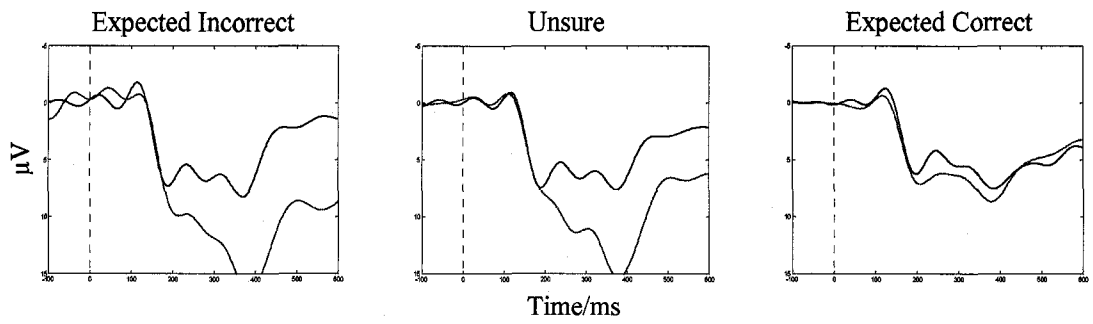


Figure 3.11 Experiment 2 ERN for outcomes of different expectancy.

Figure 3.11 seems to suggest that the negativity following error feedback is the same regardless of the participants' expectation and, indeed, there is no significant difference in the peak-to-peak ERN measures between the different confidence conditions ($F(2,24) = 0.112$, $MSE = 2.734$, $p = 0.867$). The major difference between the confidence conditions seems to involve the waveforms following correct feedback. There appeared to be a much larger P300 following correct feedback when participants were unsure about their response or thought they were going to be presented with incorrect feedback. This, however, fits with what is known about the functional significance of the P300, which has consistently been found to be larger following unexpected events. However, if the P300 is a component of the error waveform as well as the correct waveform the peak-to-peak measures will be contaminated by the different P300 contributions in the different error conditions. This issue will be addressed later in the discussion of frequency effects on the ERN. While the ERN does not appear to differ between the confidence ratings this could again be masking effects related to learning. Figure 3.12 shows the error and correct waveforms for the different confidence ratings for the first third and second two thirds of the block.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

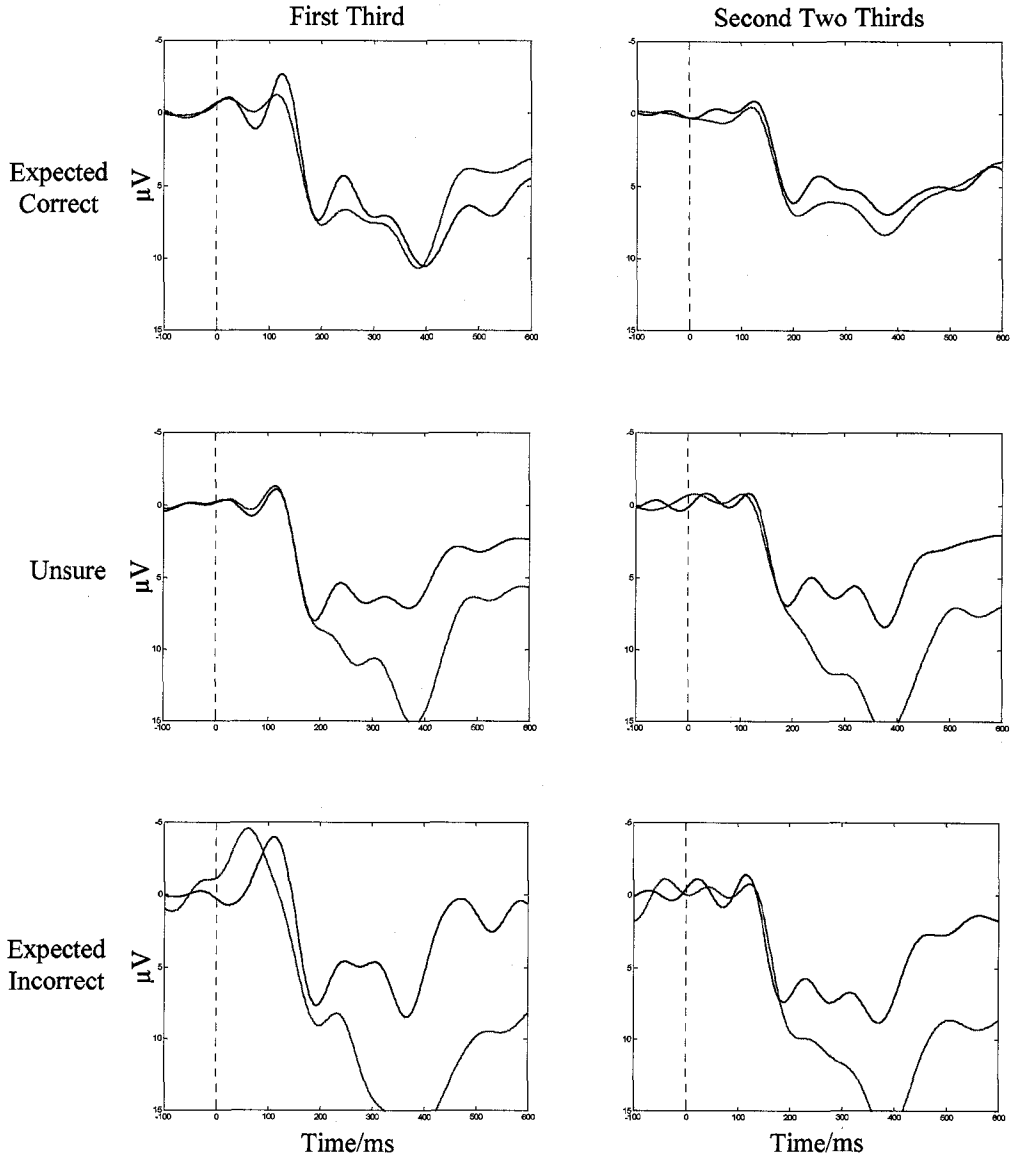


Figure 3.12 Experiment 2 change in ERN over the block for the different confidence ratings.

Again, these waveforms suggest that the ERN decreases over the block. However, while none of the differences in the error waveforms between the first and second parts of the block were significantly different, the peak-to-peak measure for both the expected correct and expected incorrect groups approached significance (expected correct: $F(1,12) = 4.049$, $MSE = 6.964$, $p = 0.0672$; unsure: $F(1,12) = 0.723$, $MSE = 1.816$, $p = 0.412$; expected incorrect: $F(1,12) = 3.250$, $MSE = 14.25$, $p = 0.096$). Combining the two

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expected groups (expected correct and expected incorrect) yielded a significant effect ($F(1,12) = 5.435$, $MSE = 1.179$, $p = 0.038$). This suggests that the ERN decreased more in the second part of the block for trials in which the participant was confident in their choice.

Another way to investigate the relative effects of learning and confidence is to examine a condition in which the confidence judgments remain the same, but learning nevertheless takes place. The weak association condition is precisely this one, where unsure ratings are given at the early in the block because the mapping is not known and unsure ratings are given later in the block because both error and correct feedback are equally likely. Figure 3.9 shows this condition and the results presented earlier reveal the difference in the ERN is not significant.

The design of this study also allowed for comparison of error and correct waveforms following feedback of different frequencies. This is potentially important because changes in frequency result in large potentials such as the N200 and P300 that can overlap with the ERN. Therefore, some studies (e.g., Holroyd et al., 2004) have grouped waveforms by frequency and measured the difference between them. While this approach avoids the problems of overlapping different frequency related components it is also subject to others. One main problem is that a subtraction is made across conditions, so while frequency may remain constant in the two conditions other factors may be different. Another is that the effect of frequency on the overlapping components (here the N200/P300) must be the same across conditions. Figure 3.13 shows the error and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

correct waveforms following frequent, equal and infrequent feedback. The left panel shows waveforms following error feedback in the random condition (that occur 75% of the time) and correct feedback when the strong association has been learnt (that occur approximately 75% of the time). The center panel shows the error and correct waveforms during the last third of the weak condition, in which error and correct feedback is presented with approximately the same frequency (approximately 50%). The right panel shows waveforms following error feedback from the last three quarters of the strong association condition and correct feedback in the random condition (both of which occur approximately 25% of the time). As numerous studies have found, these plots suggest that the amplitude of the P300 increases with decreasing frequency.

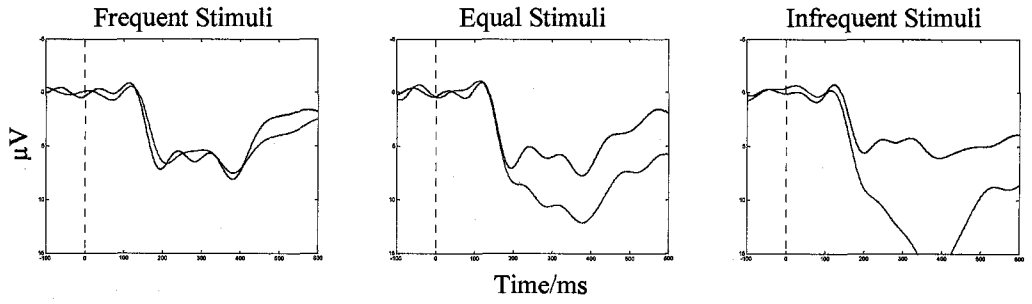


Figure 3.13 Experiment 2 ERN for stimuli of different frequency.

While there does not appear to be any difference between the error waveforms, the difference between the error and correct waveforms appears to increase for less frequent feedback. This is reflected in a significant interaction in a 2 (accuracy) x 3 (frequency) ANOVA ($F(2,24) = 6.708$, $MSE = 9.819$, $p = 0.0093$). While the effects of frequency are clear for the correct trials, the same is not the case for the error trials. If the P300 affected the error and correct waveforms equally, for the error waveform to be the same across all conditions, the difference in ERN across the conditions must be equivalent to the difference in the P300 across the condition. However, this seems unlikely because of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the size and time-course of the P300. Therefore, it appears as though this measure of the ERN is inappropriate in this case.

As in Experiment 1, errors can be divided into ignorance and violation errors (Figure 3.14). Again, the effect of the P300 is evident on correct trials. This result appears to be somewhat contrary to those found in Experiment 1 that appeared to suggest a difference between ignorance and violation errors because there is, in fact, no significant difference between the two types of error using peak-to-peak analysis ($F(1,12) = 1.039$, $MSE = 24.766$, $p = 0.328$).

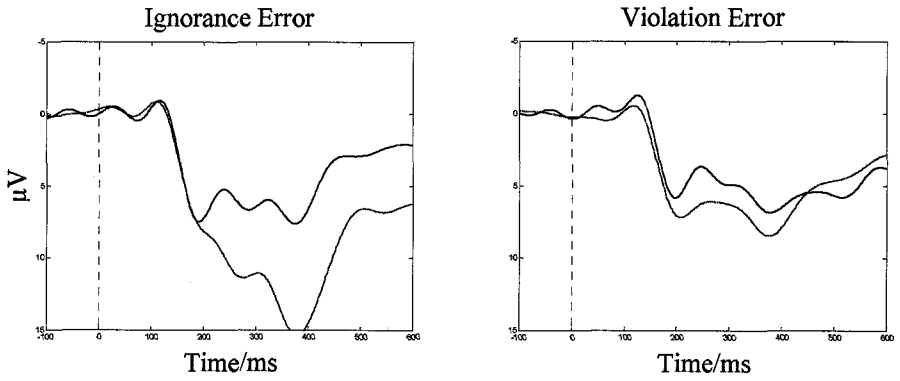


Figure 3.14 Experiment 2 ignorance and violation ERNs.

Finally, the electrophysiological effects of gaining and losing money can be investigated. Figure 3.15 shows the waveforms following value presentation for gains and losses (i.e., on correct and errors trials). No ERN-like negativity is observed following either stimulus.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

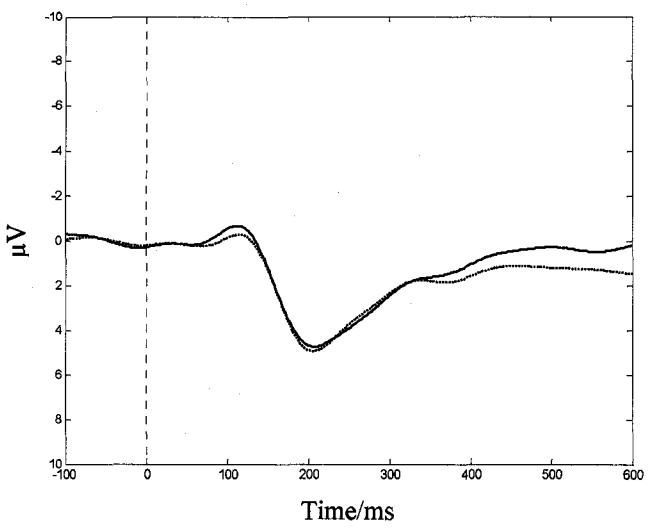


Figure 3.15 Experiment 2 waveforms following value presentation. Waveforms for gain trials are plotted in the solid line; losses are plotted in the dotted line.

Because of the experimental design the effects of gaining or losing more or less money than expected can also be investigated. Participants were presented with the value 25 twice as frequently as either 5 or 50. Therefore, there were two types of outcome that were worse than expected, a gain of 5 (less than the expected gain of 25) and a loss of 50 (more than the expected loss of 25) and two types of outcome that were better than expected, a gain of 50 and a loss of 5. These types of better and worse than expected outcomes have been combined and are shown in Figure 3.16 compared with the waveforms for expected outcomes. As in Experiment 1, no ERN-like component is observed following any of the outcomes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

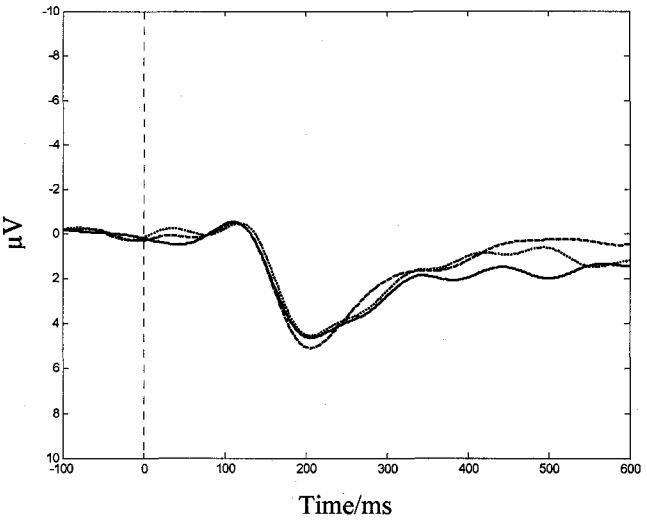


Figure 3.16 Experiment 2 waveforms for expected (dashed line), better than expected (solid line) and worse than expected values (dotted line).

The waveforms in both Experiments 1b and 2 were different from those observed in Experiment 1a in that there appeared to be a second negativity following the ERN. This has been noted informally in a number of experiments, but has rarely been the direct subject of investigation. One possibility is that subjects do, indeed, show another negativity following the ERN. Luu and Tucker (2001) have suggested this might be a result of the on-going theta oscillation and Fiehler et al. (submitted) noted a similar negativity following corrected errors (although this seems unlikely to be the case in these experiments since corrections could not be made). Another possible explanation could be that some subjects show the ERN later than during the usual 200-300 ms time window. When averaged together this might result in a waveform that appears to show two negativities, whereas in actuality there is only one present in the waveform for each individual subject. Figure 3.17 shows 4 subjects from Experiment 1b. The upper left panel shows a subject with a large ERN between 200 and 300 ms. While in this case there was a second negativity, it did not peak until about 450 ms (well after the second

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

negativity in the averaged waveform). The upper right panel shows a subject with a later and slower ERN that could potentially contribute to the second negativity in the averaged waveform. However, the two lower panels show waveforms from subjects that distinctly show a second negativity.

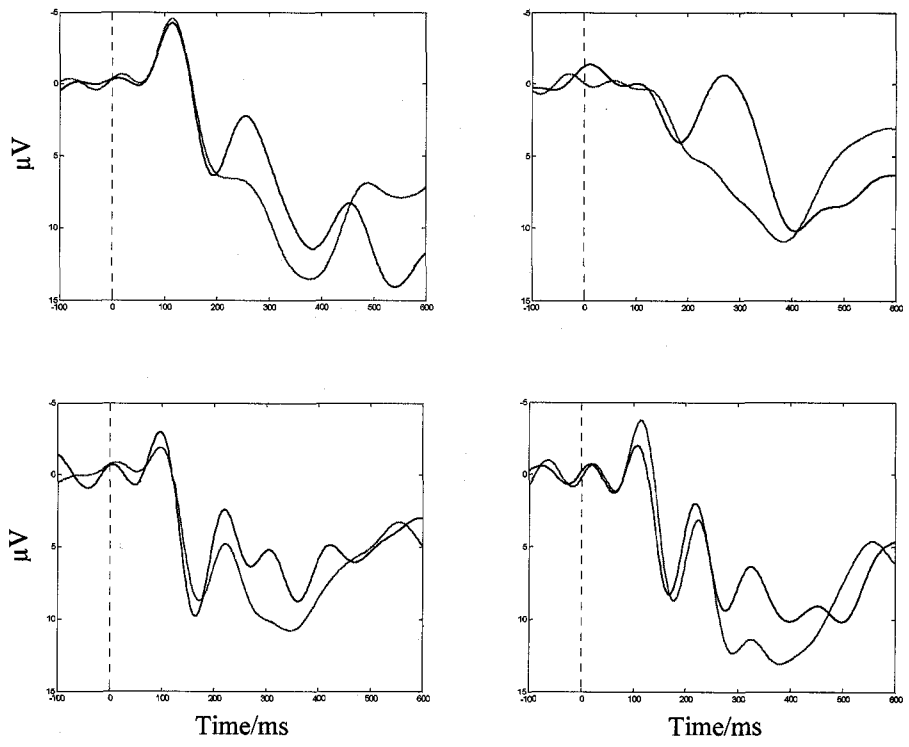


Figure 3.17 Error and correct waveforms from four individual subjects in Experiment 1b.

Figure 3.18 displays similar waveforms for two subjects in Experiment 2. Again, while there were subjects that showed an ERN with no following negativity, there was also a subset of subjects that showed the second negativity.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

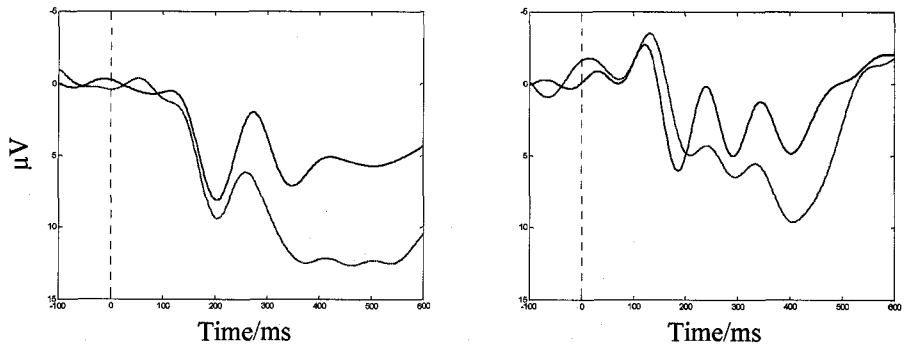


Figure 3.18 Error and correct waveforms from two individual subjects in Experiment 2.

This suggests that the existence of a second negativity might be a real phenomenon. However, what its functional significance might be and why it only manifests in a subgroup of subjects is still unknown and perhaps warrants further investigation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Experiment 2 Discussion

The behavioral results were consistent with those found in Experiment 1. Stronger association strengths were learnt more quickly and confidence increased as the association was learnt (at least for the strong and random conditions). Furthermore, participants were more confident in their correct responses than their incorrect ones.

The basic electrophysiological results also replicated those in Experiment 1. An ERN was observed following error feedback, although it had a slightly more frontal scalp distribution. By examining the learning curves for the different association strengths, Experiment 2 divided the block into parts that more accurately reflected learning progression than Experiment 1. This revealed that the ERN was reduced later in the block, when the probability mapping had been learnt and that this effect was driven mainly by a reduction in the ERN amplitude in the strong association condition. Another result that was consistent with the first experiment was that no effect of value was observed, nor was any effect seen following rewards or penalties that were better or worse than expected. Other findings, however, were not consistent with Experiment 1. The ERN was not found to decrease with increasing confidence for all association strengths, although it did decrease over the course of the block for high confidence ratings. However, this suggests that ERN amplitude is more dependent on how well the association has been learnt than on participants' confidence itself, which is more consistent with the pattern seen in Experiment 1b than 1a. Another inconsistency between the two experiments was the lack of a difference between ignorance and violation errors in Experiment 2. Nevertheless, there was no indication of an increase in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ERN amplitude following violation errors that would be predicted by the RL-ERN. One potential reason for these inconsistencies might be that in Experiment 2, participants only had to learn one association at a time. Having three associations to learn at the same time might have led to less attention being paid to an association that was learnt more quickly than the others, which could potentially have resulted in a further decrease in ERN amplitude. This might have especially had an effect in Experiment 1a because one probability mapping was very strong (85%) and could have been learnt very quickly. Another effect of having to learn only one association at a time is that participants may have held clearer expectancies and have tried to constantly refine their behavior rather than concentrate on the associations they had yet to learn. Also, the change in confidence rating system, while affording some benefits, reduced the discrimination between different confidence levels, which may have contributed to the differences in the results relating to confidence, although much of the analysis in Experiment 1 concentrated on grouped confidence ratings similar to those in Experiment 2.

The observation that the ERN decreases over the course of the block is again inconsistent with the RL-ERN theory. This theory predicts that the ERN is larger the worse the actual outcome is compared with the predicted outcome. This would be the case when the association is learnt and a strong expectation is held; so according to the RL-ERN, the ERN should be largest towards the end of learning. While a previous study (Holroyd and Coles, 2002) found that the ERN decreased over the course of the block, this was only observed for stimuli that had a 100% mapping (i.e., stimuli were always mapped to the same response). These later error trials were associated with a larger response-locked

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ERN for which the explanation was that the information about the accuracy of the response was available immediately (as in the case of the traditional response ERN). This cannot be used an as explanation for these results as in neither study was a response-locked ERN observed nor was accuracy information immediately available. Despite this result, the finding of an increased difference between the error and correct waveforms when the feedback is matched for frequency is consistent with the RL-ERN and is analogous to results found by Holroyd et al. (2004) who used reward and penalty feedback rather than accuracy feedback. However, in Experiment 2, reward and accuracy feedback were somewhat confounded because all errors were penalized and all correct responses were rewarded. However, there are a number of caveats to this measure of the ERN (described earlier) and these results are only consistent with the RL-ERN if it is indeed the case that the P300 overlap affects the waveforms following error and correct feedback equivalently.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter IV

General Discussion

The Basic Results

The general pattern across the experiments presented here is reasonably clear. The ERN decreases in amplitude over the course of learning. Experiment 2 suggests this is certainly the case for strong associations. This observation is compatible with Experiment 1a, in which the overall effect could have been driven by the strong (85%) condition. What is less clear is whether this decrease is more strongly related to confidence in the particular response or how well the association has been learnt. While Experiment 1a seems to suggest that the decrease might be more strongly related to confidence both Experiments 1b and 2 appear to provide evidence that it is more closely related to how well the association is learnt, since the ERN was found to decrease over the block on high confidence trials. In none of the experiments did the outcome value affect the ERP. This is quite surprising since a number of studies have shown effects of winning and losing money on the ERN/MFN (e.g., Gehring and Willoughby, 2002; Nieuwenhuis et al., 2004; Holroyd et al., 2004). A potential explanation for this might be that the participants paid less attention to or cared lass about the values being presented because they were more involved in the learning task.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Implications for Theories of the ERN

The RL-ERN received mixed support from these experiments. While Experiment 2 resulted in some findings consistent with the RL-ERN (if the ERN is measured matched for frequency), it also uncovered some that are inconsistent with it. The RL-ERN has trouble explaining the results that showed the ERN decreased in size over learning. The RL-ERN theory would predict that as learning takes place, the expectancy for the most common event would increase (correct feedback for strong associations) and expectancy for the least common event (error feedback for the strong associations) would decrease. Therefore, when error feedback is presented late in learning, the difference between the actual and predicted events would be greater than when it is presented early in learning. This would especially be the case for strong associations that are learnt quickly. Therefore, it is notable that Experiment 2 found the decrease to be significant only for the strongest association strength and in Experiment 1a, in which there was a significant decrease, a very strong association was presented most frequently.

The results of these experiments can also be considered in terms of the other ERN theories. The conflict-monitoring theory is perhaps the simplest to consider. It cannot easily explain any feedback-related negativities because they occur after the response is made, when there is no response conflict to be detected. Therefore, to the extent that the original response ERN (for which the model was designed to explain) and feedback ERN or MFN reflect the same psychological process, the observation of any feedback negativity is a major problem for the conflict-monitoring theory. One potential way to reconcile the conflict-monitoring theory with the observation of feedback negativities

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

would be to postulate that the conflict-monitoring system is sensitive to other types of conflict as well as response conflict (e.g., stimulus conflict).

The results presented here are also problematic for the error detection theory of the ERN. Different versions of this theory propose that the ERN reflects the comparison between the correct and error representations, the output of that comparison process or a combination of the two processes (e.g., Falkenstein et al., 2001). However, these processes would be the same regardless of learning. The comparison process is comparing the same representations at all times during learning so the degree to which the comparison system is engaged to detect the mismatch between them is constant throughout learning. Similarly, the output of the comparison would also be the same because the difference between the representations remains constant. Therefore, it is difficult to see how the error detection theory as it stands could account for a change in ERN amplitude over learning without incorporating expectation into the theory, the nature of which would have to be fully specified.

The Affective Theory

In contrast to the prediction of the RL-ERN theory, the results seem to suggest that the ERN is larger earlier in learning. This is consistent with the animal and neuroimaging studies in which ACC activity has been observed only in the early stages of learning (e.g., Erickson et al., 2004; Gemba et al., 1986). This result potentially gives some insight into the nature of the representation of the probabilistic association being learnt. Reinforcement learning theories propose that an expectancy is derived from past events

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and current events update that expectancy. However, the results here suggest a slightly different conception of expectancy that includes a representation of the distribution of events. Once a probability distribution is learnt, infrequent events are not necessarily unexpected, as long as they occur with the same frequency as the learnt distribution. However, this kind of representation of the distribution of events has not been incorporated into a reinforcement learning framework. Furthermore, the meaning of expectation and the underlying structure of the representation of the association being learnt need to be specified. Investigation into the effects of infrequent expected (e.g., in a predictable sequence) versus unexpected (e.g., in an unpredictable sequence) stimuli might be able to elucidate the nature of this representation. Nevertheless (notwithstanding the theory of Berridge and Robinson, 1998), there is still considerable evidence that the mesencephalic dopamine system is involved in processing reward and punishment information and that this system is connected with the ACC (e.g., Holroyd and Coles, 2002), therefore, it might prove unwise to discount its involvement in traditional reinforcement learning completely.

The affective theory is the least well specified of the ERN theories and perhaps has not yet earned the title 'theory.' While the seed of the idea has been around for a long time, it has not yet germinated into anything much more substantial and as will be seen, it might also be better described as a motivational theory rather than an affective one. At its simplest, it predicts that the ERN would be larger for affectively 'worse' errors. This then raises the questions of what would constitute an affectively 'worse' error and (to account for the results presented here) whether or not errors earlier in learning are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

'worse' than those later of learning. In order to tackle these questions it is helpful first to consider emotion's role in general.

It appears that the emotional system monitors for motivationally significant events. When it encounters such an event a number of systems are engaged in order to deal effectively with it. One of those systems is the attentional system; when we encounter a motivationally significant event it captures and holds our attention presumably, at least partly, to enhance psychological processing of it. Not only do emotional events attract our attention they also cause physiological arousal (that can be observed in, for example, heart rate and galvanic skin response changes). This arousal generally prepares us to engage in some physical (or possibly psychological) action related to the emotional event. Furthermore, a learning system is likely to be activated to record the circumstances under which the event took place. So perhaps, then, what is meant by 'affectively worse' is actually 'motivationally more significant.' However, these are usually negative events because they are the ones that are likely to motivate some action to avoid or escape the event; there would be little reason to motivate a change in behavior on the occurrence of a good, or positive, event (although learning under what conditions it occurred might be useful). Now we can translate the question into whether the feedback early in learning is motivationally more significant than later in learning. The answer seems to be that it is. Error feedback early in learning motivates a change in behavior, that of pressing a different button next time and attempting to learn the mapping. Error feedback later in learning though does not motivate this change in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

behavior or learning because the mapping has been learnt and the most appropriate response has been made (nor does it provide further information about the mapping).

This proposition for the functional significance of the ERN fits well with the putative functions of the ACC. The ACC has been shown to be active in tasks requiring attention (e.g., Peterson et al., 1999) and has close links to the dorso-lateral prefrontal cortex (e.g., Gehring and Knight, 2000; Ullsperger et al., 2002), which is thought to have a role in cognitive control and inhibition (e.g., Botvinick et al., 2001). The ACC has frequently been shown to be active in emotional tasks (e.g., Phan et al., 2002) and to be related to the control of autonomic functions (e.g., Critchley et al., 2003). Furthermore, the cingulate motor areas are closely connected with the premotor and supplementary motor regions suggesting a role in preparation of action. Also, as the RL-ERN theory proposes, the ACC is linked to the reward and punishment systems.

It has also been claimed that the ACC contains the representations for motivational drives and pain, which are likely to motivate behavioral responses when active (Sewards and Sewards, 2003). Furthermore, the ACC is active on presentation of motivationally significant stimuli that themselves are not emotional in nature, but motivate a response (Fichenholtz et al., 2004). This comparison may highlight an important distinction between motivationally significant stimuli. Some, that are critical for everyday survival, could be hard-wired and are probably permanent, whereas others can be learnt flexibly, depending on the task demands. This, of course, makes good evolutionary sense. It would be maladaptive not to have access to the motivational significance of some stimuli

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

straight away, but it would be equally maladaptive not to be able to learn what stimuli are and are not motivationally significant in a given environment.

Another issue is to what extent this theory of the ERN complements the existing theories (and how it differs from them) and how it explains other experimental findings. The response ERN occurs when people realize they have committed an error. This is a motivationally significant event that engages attention and strategic control mechanisms that affect future performance, usually manifesting in immediate remedial behavior and post-error slowing. Motivational instructions either to emphasize speed or accuracy (e.g., Gehring et al., 1993) or to correct mistakes as soon as possible (e.g., Fiehler et al, submitted) affect the amplitude of the ERN. Furthermore, it has been suggested that the ERN occurs after the error has been detected (e.g., Gehring and Fencsik, 1999; Rodriguez-Fornells et al., 2002) which might indicate its time-course is compatible with an evaluative function. This affective account of the ERN is also consistent with the enhanced ERN found in OCD patients and highly anxious experimental participants. Overactivity of a evaluation and motivation system might result in over-engagement of attentional and action systems that might lead to obsessions and compulsive behaviors to alleviate them.

This affective account of the ERN appears to fit less well with the conflict-monitoring theory. However, an effort can be made to understand how the conflict-monitoring theory might fit into this framework. The main evidence for the conflict-monitoring theory comes from ACC activity (frequently dorsal), measured by fMRI, on correct trials.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Electrophysiological evidence seems to suggest that the ERN is not sensitive to conflict on error trials (e.g., Luu et al., 2000), although there is some evidence of an electrophysiological marker of conflict on correct trials, the N2 (e.g., van Veen and Carter, 2002). However, it has also been suggested that the dorsal medial-frontal cortex, that is, the premotor and supplementary motor areas (and perhaps the cingulate motor areas) are involved in sending a 'command volley' to the primary motor cortex to execute movement (e.g., Gabriel et al., 2002). If multiple commands are processed or sent to primary motor cortex by these premotor areas, more activity might be expected. Therefore, the activity in these areas might not necessarily index response conflict between mutually incompatible response representations encoded elsewhere, but could reflect the processing or sending of multiple 'command volleys' to primary motor cortex.

The RL-ERN theory is, in some respects, quite similar to the proposition made here and goes into explicit detail about how learning is carried out following events that are worse than expected. In general, events that are worse than expected are motivationally significant; therefore, an ERN will frequently be observed following them. However, as this series of experiments suggests, they need not necessarily be motivationally significant and therefore are not necessarily followed by an ERN. It is also necessary to be explicit concerning what and how expectation is being measured (e.g., the pure reinforcement learning expectancy or one subject to other cognitive influences). The RL-ERN theory also specifies how learning about worse than expected events might occur and, indeed, this might be part of a larger affective system. However, while reinforcement learning theory is a popular way of explaining many empirical findings in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the learning literature it is not universally accepted (e.g., Berridge and Robinson, 1998). This suggests there may be additional ways to about motivationally significant events.

One key study that implicated medial-frontal ERN-like activity (the MFN) in evaluation of motivation significance showed it was sensitive to monetary gains and losses in a gambling task rather than accuracy (Gehring and Willoughby, 2002). A further finding was that modulation of behavior on trials following losses was related to ERN amplitude, again suggesting a connection between the ERN and motivation to change behavior. The interpretation of the feedback ERN given here suggests that the ERN and MFN might reflect at least somewhat similar psychological processes and, therefore, be a parsimonious explanation of them both.

If this analysis of the functional significance of the ERN approaches the truth, outstanding questions remain concerning what particular feature (or combination of features) of this system the ERN reflects. Potentially, it could be the evaluative, attentional, arousal, action or learning component, or a modulatory component that processes multiple types of information. These questions can and are being approached though careful experimentation involving numerous different manipulations. For example, Rodriguez-Fornells and Munte (2004) have shown that the ERN does not appear to be related to the number of simultaneously active motor programs suggesting its primary role might not be in motor preparation. Also, van der Veen et al. (2004) have shown that the ERN is not directly correlated with the change in cardiac response following an error, suggesting it is not directly involved in an arousal mechanism. This is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

not to say, though, that the ACC in general is not involved in these functions, merely that the ERN itself does not appear to reflect these psychological functions. Indeed, the cortical proximity of brain regions involved in these processes assessed by other investigative techniques suggests a strong functional and physiological relationship between them. Furthermore, theoretical frameworks and explicit models such as the error detection, conflict-monitoring and RL-ERN models are also invaluable in specifying predictions and allowing them to be tested in novel and sophisticated ways.

Conclusion

The increasing interest in the functional significance of the ERN reflects a deeper interest in understanding the monitoring and control processes of the human mind. The emotional system monitors and evaluates on-going events and motivates us to behave in adaptive ways. There is increasing evidence that the ERN plays an evaluative and motivational role and may be involved in a wider emotional system. However, owing to the number of theories attempting to explain it, the functional significance of the ERN is still controversial, which can only motivate more research and result in a better understanding of both theoretical issues and empirical evidence.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bibliography

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Aggleton, J. P. (1992). The functional effects of amygdala lesions in humans: A comparison with findings from monkeys. In J. P. Aggleton (Ed.) Amygdala: Neurobiological aspects of emotion, memory, and mental dysfunction. (pp. 485-503). New York, NY, USA: Wiley-Liss.

Aggleton, J. P. (1993). The contribution of the amygdala to normal and abnormal emotional states. Trends in Neurosciences, 16(8), 328-333.

Alain, C., McNeely, H. E., He, Y., Christensen, B. K., & West, R. (2002). Neurophysiological evidence of error-monitoring deficits in patients with schizophrenia. Cerebral Cortex, 12(8), 840-846.

Allman, J. M., Hakeem, A., Erwin, J. M., Nimchinsky, E., & Hof, P. (2001). The anterior cingulate cortex - the evolution of an interface between emotion and cognition. Annals of the New York Academy of Sciences, 935, 107-117.

Attneave, F. (1957). Physical determinants of the judged complexity of shapes. Journal of Experimental Psychology, 53(4), 221-227.

Badgaiyan, R. D., & Posner, M. I. (1998). Mapping the cingulate cortex in response selection and monitoring. Neuroimage, 7(3), 255-260.

Bancaud, J., & Talairach, J. (1992). Clinical semiology of frontal lobe seizures. Advances in Neurology, 57, 3-58.

Band, G. P. H., & Kok, A. (2000). Age effects on response monitoring in a mental-rotation task. Biological Psychology, 51(2-3), 201-221.

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. Journal of Cognitive Neuroscience, 12(2), 298-309.

Barrett, J., Pike, G. B., & Paus, T. (2004). The role of the anterior cingulate cortex in pitch variation during sad affect. European Journal of Neuroscience, 19(2), 458-464.

Barris, R. W., & Schuman, H. R. (1953). Bilateral anterior cingulate gyrus lesions. Syndrome of the anterior cingulate gyri. Neurology, 3, 44-52.

Bates, A. T., Kiehl, K. A., Laurens, K. R., & Liddle, P. F. (2002). Error-related negativity and correct response negativity in schizophrenia. Clinical Neurophysiology, 113(9), 1454-1463.

Bates, A. T., Liddle, P. F., Kiehl, K. A., & Ngan, E. T. C. (2004). State dependent changes in error monitoring in schizophrenia. Journal of Psychiatric Research, 38(3), 347-356.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Benedict, R. H. B., Shucard, D. W., Maria, M. P. S., Shucard, J. L., Abara, J. P., Coad, M. L., Wack, D., Sawusch, J., & Lockwood, A. (2002). Covert auditory attention generates activation in the rostral/dorsal anterior cingulate cortex. Journal of Cognitive Neuroscience, 14(4), 637-645.

Berlyne, D. E. (1957). Conflict and choice time. British Journal of Psychology, 48(2), 106-118.

Bernstein, P. S., Scheffers, M. K., & Coles, M. G. H. (1995). "Where did I go wrong?" A psychophysiological analysis of error detection. Journal of Experimental Psychology: Human Perception & Performance, 21(6), 1312-1322.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? Brain Research Reviews, 28(3), 309-369.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. Psychological Review, 108.

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of frequency, inhibition and errors. Cerebral Cortex, 11(9), 825-836.

Brazdil, M., Roman, R., Falkenstein, M., Daniel, P., Jurak, P., & Rektor, I. (2002). Error processing - Evidence from intracerebral ERP recordings. Experimental Brain Research, 146(4), 460-466.

Brazier, M. A. B. (1984). A history of neurophysiology in the 17th and 18th. New York, NY, USA: Raven Press.

Brazier, M. A. B. (1988). A history of neurophysiology in the 19th century. New York, NY, USA: Raven Press.

Burack, J. A., Charman, T., Yirmiya, N., & Zelazo, P. R. (2001). Development and autism: Messages from developmental psychopathology. J. A. Burack, & T. Charman Development of autism: Perspectives from theory and research (pp. 3-15). Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Burns, S. M., & Wyss, J. M. (1985). The involvement of the anterior cingulate cortex in blood-pressure control. Brain Research, 340(1), 71-77.

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

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging: Validation study with functional MRI. Human Brain Mapping, 6(4), 270-282.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Butterfield, B., & Mangels, J. A. (2003). Neural correlates of error detection and correction in a semantic retrieval task. Cognitive Brain Research, 17(3), 793-817.

Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. Neuroscience and Biobehavioral Reviews, 26 (3), 321-352.

Carrie, J. R. (1972). Applications of digital general-purpose computers in a clinical EEG laboratory. American Journal of EEG Technology, Vol. 12(3), 135-140.

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(5364), 747-749.

Carter, C. S., Macdonald, A. M., Botvinick, M. M., Ross, L. L., Stenger, V. A., Noll, D., & Cohen, J. D. (2000). Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. Proceedings of the National Academy of Sciences of the United States of America, 97(4), 1944-1948.

Casey, B. J., Thomas, K. M., Welsh, T. F., Badgaiyan, R., Jennings, J. R., & Crone, E. A. (1999). Involvement of the anterior cingulate and related circuitry in inhibition of attention to competing peripheral stimuli. Journal of Cognitive Neuroscience, 57.

Cohen, R. A., Paul, R., Zawacki, T. M., Moser, D. J., Sweet, L., & Wilkinson, H. (2001). Emotional and personality changes following cingulotomy. Emotion, 1(1), 38-50.

Coles, M. G. H., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. Biological Psychology, 56(3), 173-189.

Critchley, H. D., Mathias, C. J., Josephs, O., O'doherty, J., Zanini, S., Dewar, B. K., Cipolotti, L., Shallice, T., & Dolan, R. J. (2003). Human cingulate cortex and autonomic control: Converging neuroimaging and clinical evidence. Brain, 126, 2139-2152.

Crone, E. A., van der Veen, F. M., van der Molen, M. W., Somsen, R. J. M., van Beek, B., & Jennings, J. R. (2003). Cardiac concomitants of feedback processing. Biological Psychology, 64(1), 143-156.

Danckert, J., Maruff, P., Ymer, C., Kinsella, G., Yucel, M., de Graaff, S., & Currie, J. (2000). Goal-directed selective attention and response competition monitoring: Evidence from unilateral parietal and anterior cingulate lesions. Neuropsychology, 14(1), 16-28.

Davidson, R. J., Abercrombie, H., Nitschke, J. B., & Putnam, K. (1999). Regional Brain Function, Emotion and Disorders of Emotion. Current Opinion in Neurobiology, 9(2), 228-234.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. Molecular Psychiatry, 6(1), 13-34.

De Molina, A. F., & Hunsperger, R. W. (1959). Central representation of affective reactions in forebrain and brain stem: Electrical stimulation of amygdala, stria terminalis, and adjacent structures. Journal of Physiology, 145, 251-265.

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

Devinsky, O., & Morrell, M. J. (1995). Contributions of anterior cingulate cortex to behaviour. Brain, 118(1), 279-306.

Dikman, Z. V., & Allen, J. J. B. (2000). Error monitoring during reward and avoidance learning in high- and low-socialized individuals. Psychophysiology, 37(1), 43-54.

Drevets, W. C., & Raichle, M. E. (1998). Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: Implications for interactions between emotion and cognition. Cognition & Emotion, 12(3), 353-385.

Ebert, D., & Ebmeier, K. P. (1996). The role of the cingulate gyrus in depression: From functional anatomy to neurochemistry. Biological Psychiatry, 39(12), 1044-1050.

Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. Science, 302(5643), 290-292.

Erickson, K. I., Milham, M. P., Colcombe, S. J., Kramer, A. F., Banich, M. T., Webb, A., & Cohen, N. J. (2004). Behavioral conflict, anterior cingulate cortex, and experiment duration: Implications of diverging data. Human Brain Mapping, 21(2), 98-107.

Estrin, T. (1965). On-line electroencephalographic digital computing system. Electroencephalography and Clinical Neurophysiology, 19(5), 524-526.

Falkenstein, M. (2004). ERP correlates of erroneous performance. In M. Ullsperger, & M. Falkenstein (Eds.) Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring. Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.

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 & Clinical Neurophysiology, 78(6), 447-455.

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

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in go/no-go tasks and their relation to inhibition. Acta Psychologica, 101(2-3), 267-291.

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (2001). Changes of error-related ERPs with age. Experimental Brain Research, 138(2), 258-262.

Falkenstein, M., Hoormann, J., Hohnsbein, J., & Hielscher, H. (2002). Event-related potentials after motor errors in patients with Parkinson's disease. Journal of Psychophysiology, 16(4), 228.

Fallgatter, A. J., Bartsch, A. J., Zielasek, J., & Herrmann, M. J. (2003). Brain electrical dysfunction of the anterior cngulate in schizophrenic patients. Psychiatry Research - Neuroimaging, 124(1), 37-48.

Fassbender, C., Murphy, K., Foxe, J. J., Wylie, G. R., Javitt, D. C., Robertson, I. H., & Garavan, H. (2004). A topography of executive functions and their interactions revealed by functional magnetic resonance imaging. Cognitive Brain Research, 20(2), 132-143.

Faymonville, M.-E., Roediger, L., Del Fiore, G., Delgueldre, C., Phillips, C., Lamy, M., Luxen, A., Maquet, P., & Laureys, S. (2003). Increased cerebral functional connectivity underlying the antinociceptive effects of hypnosis. Cognitive Brain Research, 17(2), 255-262.

Fichtenholtz, H. M., Dean, H. L., Dillon, D. G., Yamasaki, H., Mccarthy, G., & Labar, K. S. (2004). Emotion-attention network interactions during a visual oddball task. Cognitive Brain Research, 20(1), 67-80.

Fiehler, K., Ullsperger, M., & von Cramon, D. Y. (submitted). Electrophysiological correlates of error correction.

Gabriel, M., Burhans, L., & Talk A. (2002). The cingulate cortex. In V. S. Ramachandran (Ed.) Encyclopedia of the Human Brain . Academic Press .

Garavan, H., Ross, T. J., Murphy, K., Roche, R. A., & Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior: Inhibition, error detection, and correction. Neuroimage, 17(4), 1820-9.

Gehring, W. J., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1995). A brain potential manifestation of error-related processing. In G. Karmos, M. Molnar, V. Csepe, I. Czigler, & J. E. Desmedt (Eds.) Perspectives of Event-Related Potentials Research, Journal of Electroencephalography and Clinical Neurophysiology, Supplement 44 (pp. 287-296).

Gehring, W. J., & Fencsik, D. E. (1999). Slamming on the brakes: An electrophysiological study of error response inhibition. Journal of Cognitive Neuroscience, Supplement: Suppl. S, 60.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

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

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

Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. Psychological Science, 11(1), 1-6.

Gehring, W. J., & Knight, R. T. (2000). Prefrontal-cingulate interactions in action monitoring. Nature Neuroscience, 3(5), 516-520.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. Science, 295(5563), 2279-2282.

Gehring, W. J., & Willoughby, A. R. (2004). Are all medial frontal negativities created equal? Toward a richer empirical basis for theories of action monitoring. In M. Ullsperger, & M. Falkenstein (Eds.) Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring . Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.

Gemba, H., Sasaki, K., & Brooks, V. B. (1986). 'Error' potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. Neurosci Letters, 70(2), 223-7.

Giannitrapani, D., Rast, V. T., & Shulhafe, B. J. (1971). Multiple channel direct digital recording of EEG data. Behavioral Science, 16(3), 239-&.

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

Grimm, C., Schreiber, A., Kristeva-Feige, R., Mergner, T., Hennig, J., & Lucking, C. H. (1998). A comparison between electric source localisation and fMRI during somatosensory stimulation. Electroencephalography and Clinical Neurophysiology, 106(1), 22-29.

Hajcak, G., McDonald, N., & Simons, R. F. (2003). Anxiety and error-related brain activity. Biological Psychology, 64(1), 77-90.

Hajcak, G., & Simons, R. F. (2002). Error-related brain activity in obsessive-compulsive undergraduates. Psychiatry Research, 110(1), 63-72.

Harmer, C. J., Thilo, K. V., Rothwell, J. C., & Goodwin, G. M. (2001). Transcranial magnetic stimulation of medial-frontal cortex impairs the processing of angry facial expressions. Nature Neuroscience, 4(1), 17-18.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Holroyd, C. B. (2004). A note on the oddball N200 and the feedback ERN. In M. Ullsperger, & M. Falkenstein (Eds.) Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring . Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.

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.

Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the event-related brain potential associated with reward and punishment. Psychophysiology, 41(2), 245-253.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G. H., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. Nature Neuroscience, 7(5), 497-498.

Holroyd, C. B., Praamstra, P., Plat, E., & Coles, M. G. H. (2002). Spared error-related potentials in mild to moderate Parkinson's disease. Neuropsychologia, 40(12), 2116-2124.

Holroyd, C. B., & Yeung, N. (2003). Alcohol and error processing. Trends in Neurosciences, 26(8), 402-404.

Horst, R. L., Johnson, R., & Donchin, E. (1980). Event-related brain potentials and subjective probability in a learning task. Memory and Cognition, 8(5), 476-488.

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science, 302(5642), 120-122.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. Electroencephalography and Clinical Neurophysiology, 10, 370-375.

Jodo, E., & Kayama, Y. (1992). Relation of a negative ERP component to response-inhibition in a go/no-go task. Electroencephalography and Clinical Neurophysiology, 82(6), 477-482.

Johannes, S., Wieringa, B. M., Nager, W., Dengler, R., & Muente, T. F. (2001). Oxazepam alters action monitoring. Psychopharmacology, 155(1), 100-106.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Johannes, S., Wieringa, B. M., Nager, W., Rada, D., Dengler, R., Emrich, H. M., Muente, T. F., & Dietrich, D. E. (2001). Discrepant target detection and action monitoring in obsessive-compulsive disorder. Psychiatry Research: Neuroimaging, 108(2), 101-110.

Johannes, S., Wieringa, B. M., Nager, W., Rada, D., Müller-Vahl, K. R., Emrich, H. M., Dengler, R., Münte, T. E., & Dietrich, D. (2003). Tourette's syndrome and obsessive-compulsive disorder: Event-related brain potentials show similar mechanisms of frontal inhibition but dissimilar target evaluation processes. Behavioural Neurology, 14(1), 9-17.

Kiehl, K. A., Liddle, P. F., & Hopfinger, J. B. (2000). Error processing and the rostral anterior cingulate: An event-related fMRI study. Psychophysiology, 37(2), 216-223.

Kiehl, K. A., Smith, A. M., Hare, R. D., Mendrek, A., Forster, B. B., Brink, J., & Liddle, P. F. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. Biological Psychiatry, 50(9), 677-684.

Kopp, B., & Rist, F. (1999). An event-related brain potential substrate of disturbed response monitoring in paranoid schizophrenic patients. Journal of Abnormal Psychology, 108(2), 337-346.

Kopp, B., & Wolff, M. (2000). Brain mechanisms of selective learning: Event-related potentials provide evidence for error-driven learning in humans. Biological Psychology, 51(2-3), 223-246.

Koyama, T., Kato, K., Tanaka, Y. Z., & Mikami, A. (2001). Anterior cingulate activity during pain-avoidance and reward tasks in monkeys. Neuroscience Research, 39(4), 421-430.

Kropotov, J. D., Crawford, H. J., & Polyakov, Y. I. (1997). Somatosensory event-related potential changes to painful stimuli during hypnotic analgesia: Anterior cingulate cortex and anterior temporal cortex intracranial recordings. International Journal of Psychophysiology, 27(1), 1-8.

Lane, R. D., Reiman, E. M., Axelrod, B., Yun, L. S., Holmes, A., & Schwartz, G. E. (1998). Neural correlates of levels of emotional awareness: Evidence of an interaction between emotion and attention in the anterior cingulate cortex. Journal of Cognitive Neuroscience, 10(4), 525-535.

Laurens, K. R., Ngan, E. T. C., Bates, A. T., Kiehl, K. A., & Liddle, P. F. (2003). Rostral anterior cingulate cortex dysfunction during error processing in schizophrenia. Brain, 126(3), 610-622.

Levin, B., & Duchowny, M. (1991). Childhood obsessive-compulsive disorder and cingulate epilepsy. Biological Psychiatry, 30 (10), 1049-1055.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Lorenz, J., Minoshima, S., & Casey, K. L. (2003). Keeping pain out of mind: The role of the dorsolateral prefrontal cortex in pain modulation. Brain, 126, 1079-1091.

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: Negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. Journal of Experimental Psychology: General, 129(1), 43-60.

Luu, P., Flaisch, T., & Tucker, D. M. (2000). Medial frontal cortex in action monitoring. Journal of Neuroscience, 20(1), 464-469.

Luu, P., & Posner, M. I. (2003). Anterior cingulate cortex regulation of sympathetic activity. Brain, 126, 2119-2120.

Luu, P., & Tucker, D. M. (2001). Regulating action: Alternating activation of midline frontal and motor cortical networks. Clinical Neurophysiology, 112(7), 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.

Mathalon, D. H., Bennett, A., Askari, N., Gray, E. M., Rosenbloom, M. J., & Ford, J. M. (2003). Response-monitoring dysfunction in aging and Alzheimer's disease: An event-related potential study. Neurobiology of Aging, 24(5), 675-685.

Mathalon, D. H., Fedor, M., Faustman, W. O., Gray, M., Askari, N., & Ford, J. M. (2002). Response-monitoring dysfunction in schizophrenia: An event-related brain potential study. Journal of Abnormal Psychology, 111(1), 22-41.

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a go/no-go response inhibition task. Human Brain Mapping, 12(3), 131-143.

Milham, M. P., Banich, M. T., Claus, E. D., & Cohen, N. J. (2003). Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. Neuroimage, 18(2), 483-493.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. Cognitive Brain Research, 12(3), 467-473.

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.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Nelson, J. K., Reuter-Lorenz, P. A., Sylvester, C. Y. C., Jonides, J., & Smith, E. E. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. Proceedings of the National Academy of Sciences of the United States of America, 100(19), 11171-11175.

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

Nieuwenhuis, S., Talsma, D., Coles, M. G. H., Holroyd, C. B., Kok, A., & van der Molen, M. W. (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. Cognitive, Affective & Behavioral Neuroscience, 2(1), 19-36.

Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2004). Stimulus modality, perceptual overlap, and the go/no-go N2. Psychophysiology, 41(1), 157-160.

Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004). Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. Cerebral Cortex, 14(7), 741-747.

Niki, H., & Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. Brain Research, 171(2), 213-224.

Ochsner, K. N., Kosslyn, S. M., Cosgrove, G. R., Cassem, E. H., Price, B. H., Nierenberg, A. A., & Rauch, S. L. (2001). Deficits in visual cognition and attention following bilateral anterior cingulotomy. Neuropsychologia, 39(3), 219-230.

Osaka, N., Osaka, M., Morishita, M., Kondo, H., & Fukuyama, H. (2004). A word expressing affective pain activates the anterior cingulate cortex in the human brain: an fMRI study. Behavioural Brain Research, 153(1), 123-127.

Pailing, P. E., & Segalowitz, S. J. (2003). The effect of uncertainty in error monitoring: An ERP investigation. Psychophysiology, 40, S65.

Park, D. C., & Gutchess, A. H. (2002). Aging, cognition, and culture: A neuroscientific perspective. Neuroscience & Biobehavioral Reviews, 26(7), 859-867.

Paus, T., Zatorre, R. J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., & Evans, A. C. (1997). Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. Journal of Cognitive Neuroscience, 9, 392-408.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain, 60, 389-443.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., & Gore, J. C. (1999). An fMRI study of Stroop Word-Color Interference: Evidence for cingulate subregions subserving multiple distributed attentional systems. Biological Psychiatry, 45(10), 1237-1258.

Phan, K. L., Liberzon, I., Welsh, R. C., Britton, J. C., & Taylor, S. F. (2003). Habituation of rostral anterior cingulate cortex to repeated emotionally salient pictures. Neuropsychopharmacology, 28(7), 1344-1350.

Phan, K. L., Taylor, S. F., Welsh, R. C., Ho, S. H., Britton, J. C., & Liberzon, I. (2004). Neural correlates of individual ratings of emotional salience: A trial-related fMRI study. Neuroimage, 21(2), 768-780.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. Neuroimage, 16(2), 331-348.

Ploner, M., Freund, H.-J., & Schnitzler, A. (1999). Pain affect without pain sensation in a patient with a postcentral lesion. Pain, 81(1), 211-214.

Ponsen, L. (1978). Data processing in EEG. American Journal of EEG Technology, 18(3), 151-158.

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science, 277(5328), 968-971.

Ridderinkhof, K. R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. H. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. Science, 298(5601), 2209-2211.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. Science, 306(5695), 443-7.

Ringler, R., Greiner, M., Kohlloeffel, L., Handwerker, H. O., & Forster, C. (2003). BOLD effects in different areas of the cerebral cortex during painful mechanical stimulation. Pain, 105(3), 445-453.

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

Rodriguez-Fornells, A., & Munte, T. F. (2004). Is the error-related negativity driven by simultaneously active motor programs? In M. Ullsperger, & M. Falkenstein (Eds.) Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring . Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Rollnik, J. D., Schroder, C., Rodriguez-Fornells, A., Kurzbuch, A. R., Dauper, J., Moller, J., & Munte, T. F. (2004). Functional lesions and human action monitoring: Combining repetitive transcranial magnetic stimulation and event-related brain potentials. Clinical Neurophysiology, 115(1), 145-153.

Ruchsow, M., Grothe, J., Spitzer, M., & Kiefer, M. (2002). Human anterior cingulate cortex is activated by negative feedback: Evidence from event-related potentials in a guessing task. Neuroscience Letters, 325(3), 203-206.

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. Journal of Experimental Psychology: Human Perception & Performance, 26(1), 141-151.

Scheffers, M. K., Humphrey, D. G., Stanny, R. R., Kramer, A. F., & Coles, M. G. H. (1999). Error-related processing during a period of extended wakefulness. Psychophysiology, 36(2), 149-157.

Schneider, F., Grodd, W., Weiss, U., Klose, U., Mayer, K. R., Nagele, T., & Gur, R. C. (1997). Functional MRI reveals left amygdala activation during emotion. Psychiatry Research-Neuroimaging, 76(2-3), 75-82.

Sewards, T. V., & Sewards, M. A. (2003). Representations of motivational drives in mesial cortex, medial thalamus, hypothalamus and midbrain. Brain Research Bulletin, 61(1), 25-49.

Shidara, M., & Richmond, B. J. (2002). Anterior cingulate: Single neuronal signals related to degree of reward expectancy. Science, 296(5573), 1709-1711.

Shima, K., & Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. 282(5392), 1335-8.

Singer, T., Seymour, B., O'doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. Science, 303(5661), 1157-1162.

Sommer, M., Meinhardt, J., & Volz, H. P. (2003). Combined measurement of event-related potentials (ERPs) and fMRI. Acta Neurobiologiae Experimentalis, 63(1), 49-53.

Stemmer, B., Segalowitz, S. J., Witzke, W., & Schonle, P. W. (2003). Error detection in patients with lesions to the medial prefrontal cortex: An ERP study. Neuropsychologia, 42(1), 118-130.

Stemmer, B., Vihla, M., & Salmelin, R. (2004). Activation of the human sensorimotor cortex during error-related processing: A magnetoencephalography study. Neuroscience Letters, 362(1), 44-47.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Swick, D., & Jovanovic, J. (2002). Anterior cingulate cortex and the Stroop task: Neuropsychological evidence for topographic specificity. Neuropsychologia, 40(8), 1240-1253.

Swick, D., & Turken, U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. Proceedings of the National Academy of Sciences of the United States of America, 99(25), 16354-16359.

Szeszko, P. R., Bilder, R. M., Lencz, T., Ashtari, M., Goldman, R. S., Reiter, G., Wu, H. W., & Lieberman, J. A. (2000). Reduced anterior cingulate gyrus volume correlates with executive dysfunction in men with first-episode schizophrenia. Schizophrenia Research, 43(2-3), 97-108.

Talairach, J., Bancaud, J., Geier, S., Bordas-Ferrer, M., Bonis, A., Szikla, G., & Rusu, M. (1973). The cingulate gyrus and human behavior. Electroencephalography and Clinical Neurophysiology, 34(1), 45-52.

Tamminga, C. A., Vogel, M., Cao, X. M., Lahti, A. C., & Holcomb, H. H. (2000). The limbic cortex in schizophrenia: Focus on the anterior cingulate. Brain Research Reviews, 31(2-3), 364-370.

Ullsperger, M., & von Cramon, D. Y. (2003). Error monitoring using external feedback: Specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. Journal of Neuroscience, 23(10), 4308-4314.

Ullsperger, M., Von Cramon, D. Y., & Muller, N. G. (2002). Interactions of focal cortical lesions with error processing: Evidence from event-related brain potentials. Neuropsychology, 16(4), 548-561.

van der Veen, F. M., Nieuwenhuis, S., Crone, E. A. M., & van der Molen, M. W. (2004). Cardiac and electro-cortical responses to performance feedback reflect different aspects of feedback processing. In M. Ullsperger, & M. Falkenstein (Eds.) Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring . Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.

Van Schie, H. T., Mars, R. B., Coles, M. G. H., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. Nature Neuroscience, 7(5), 549-554.

Van't Ent, D., & Apkarian, P. (1999). Motoric response inhibition in finger movement and saccadic eye movement: A comparative study. Clinical Neurophysiology, 110(6), 1058-1072.

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.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Vidal, F., Burle, B., Bonnet, M., Grapperon, J., & Hasbroucq, T. (2003). Error negativity on correct trials: A re-examination 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(2), 109-128.

Vitacco, D., Brandeis, D., Pascual-Marqui, R., & Martin, E. (2002). Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing. Human Brain Mapping, 17(1), 4-12.

Vogt, B. A., Berger, G. R., & Derbyshire, S. W. G. (2003). Structural and functional dichotomy of human midcingulate cortex. European Journal of Neuroscience, 18(11), 3134-3144.

Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional heterogeneity in cingulate cortex: The anterior executive and posterior evaluative regions. Cerebral Cortex, 2(6), 435-443.

Wager, T. D., Rilling, J. K., Smith, E. E., Sokolik, A., Casey, K. L., Davidson, R. J., Kosslyn, S. M., Rose, R. M., & Cohen, J. D. (2004). Placebo-induced changes in fMRI in the anticipation and experience of pain. Science, 303(5661), 1162-1167.

Walton, M. E., Bannerman, D. M., Alterescu, K., & Rushworth, M. F. S. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. Journal of Neuroscience, 23(16), 6475-6479.

Yeung, N. (2004). Relating cognitive and affective theories of the error-related negativity. In M. Ullsperger, & M. Falkenstein (Eds.) Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring. Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. Psychological Review, 111, 931-959.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.