Event-related EEG Analysis:
Simple solutions or complex computations

By

Andrew James Hanson

Doctor of Philosophy

Institute of Neuroscience

September 2016

Acknowledgements

I would first like to thank the three people that were instrumental in setting me up on this academic path, seven years ago (or nine years when you consider all of the paperwork and first drafts of PhD proposals that came before my official first day of being a student again). Mel, thank you for helping me carve out the time to keep things moving along; Kai, thank you for your constant positivity and belief in my ability and knowledge; Hamish, thank you for pushing me to keep going, helping me to change direction and persevering when everything started to look very technical.

Special thanks to Andreas for fighting my corner and helping me to explain the validity of my thoughts and equations.

…and finally many thanks to the EEG technologist for always being willing to help me find answers to my questions

Contents

Acknowledgements v

List of Tables, Figures and Equations iv

Abbreviations List viii

Abstract ix

Chapter 1. Introduction 1

1.1. Preprocessing 3

1.1.1. Filtering 3

1.1.2. 50/60 Hz ambient electrical noise 8

1.1.3. Ocular artefact removal 10

1.2. Signal Analysis Methods 19

1.2.1. Trial averaging approaches 20

1.2.2. Blind-source separation approaches 21

Chapter 2. Analysis pipeline configurations 26

2.1. Preprocessing optimization pipeline (POP) workflow 27

2.2. Preprocessing techniques and parameter selection 28

2.2.1. Cleanline plugin 29

2.2.2. Filtering 29

2.2.3. Ocular artefact removal 31

2.2.4. Baseline correction 32

2.2.5. Data pruning by epoch exclusion 33

2.2.6. Averaging 33

2.2.7. Data pruning by participant exclusion 34

2.3. POP Assessment 35

2.3.1. Signal-to-noise ratio 36

2.3.2. Signal Variance 37

2.3.3. Baseline Variance 37

2.4. Signal analysis techniques and parameter selection 38
2.4.1. Grand average informed peak detection (GA-PD) 39
2.4.2. Individual average peak detection (IAPD) 39
2.4.3. ICA informed peak detection (ICA-PD) 40
2.4.4. Component of interest peak detection (COIPD) 42
Chapter 3. Passive Auditory Task (IDAEP) 43
3.1. Background 43
3.2. Methods 46
3.2.1. Participants 46
3.2.2. Task 46
3.2.3. EEG Recording 46
3.3. Analysis 47
3.4. Results 49
3.4.1. BV exclusion criteria threshold and PVAF window limits 49
3.4.2. POP output 53
3.4.3. GA-PD 56
3.4.4. IAPD 58
3.4.5. ICA-PD 61
3.4.6. COIPD 62
3.5. Conclusions 64
Chapter 4. Passive Visual Task (LTP) 83
4.1. Background 83
4.2. Methods 86
4.2.1. Participants 86
4.2.2. Task 86
4.2.3. EEG Recording 86
4.3. Analysis 86
4.4. Results 88

4.4.1. BV exclusion criteria threshold and PVAF window limits (wild data) 88
4.4.2. POP output (pruned data) 93
4.4.3. GA-PD (pruned data) 95
4.4.4. IAPD 98
4.4.5. ICA-PD 99
4.4.6. COIPD 100
4.5. Conclusions 103
Chapter 5. Discussion 108
5.1. Conclusions 108
5.1.1. Signal analysis strategies 108
5.1.2. Preprocessing strategy optimisation and sensitivity 110
5.2. Limitations 113
5.3. Future directions 115
5.4. Summary 118
Chapter 6. Appendices 119
6.1. Supplementary Analysis: ICA algorithm training approaches 119
6.2. Supplementary Data: Median ASF statistics for the IDAEP task 127
Chapter 7. References 129

List of Tables, Figures and Equations

Table 1. POP configuration soft parameter settings..... 27
Table 2. IDAEP POP soft parameter configurations..... 48
Table 3. Wild-type POP variant signal and noise index measures and GA-PD linear model results..... 50
Table 4. Pruned POP signal and noise index measures following participant exclusion via the minimum epoch number exclusion threshold criterion approach (IDAEP task)..... 54
Table 5. Pruned POP signal and noise index measures following participant exclusion via the maximum BV exclusion threshold criterion (IDAEP task)..... 55
Table 6. Pruned GA-PD ExCrit.$Epoch_{min}$ linear and polynomial model statistics (IDAEP task)..... 56
Table 7. Pruned GA-PD ExCrit.BV_{max} linear and polynomial model statistics (IDAEP task)..... 57
Table 8. IAPD ExCrit.$Epoch_{min}$ linear and polynomial model statistics (IDAEP task)..... 59
Table 9. IAPD ExCrit.BV_{max} linear and polynomial model statistics (IDAEP task)..... 60
Table 10. ICA-PD ExCrit.$Epoch_{min}$ linear and polynomial model statistics (IDAEP task). 61
Table 11. ICA-PD ExCrit.BV_{max} linear and polynomial model statistics (IDAEP task)..... 62
Table 12. COIPD ExCrit.$Epoch_{min}$ linear and polynomial model statistics (IDAEP task).. 63
Table 13. COIPD ExCrit.BV_{max} linear and polynomial model statistics..... 63
Table 14. Gamma type POP configuration subset effect sizes for each signal analysis technique (IDAEP task)..... 69
Table 15. N1 and P2 peak latency statistics for ExCrit.BV_{max} Config.15 (IDAEP task)..... 69
Table 16. Wild POP variant preprocessing outcome measures and GA-PD analysis results..... 89
Table 17. Pruned POP variant preprocessing outcome measures for ExCrit.$Epoch_{min}$ (LTP task)..... 94
Table 18. Pruned POP variant preprocessing outcome measures for ExCrit.BV_{max} (LTP task)..... 94
Table 19. Pruned GA-PD N1b potentiation statistics (LTP task)..... 95
Table 20. IAPD N1b potentiation statistics (LTP task)..... 98
Table 21. ICA-PD N1b potentiation statistics (LTP task)..... 99
Table 22. COIPD N1b potentiation statistics (LTP task)..... 100

Table 23. POP configuration subset effect sizes (Cohen's d) for each signal analysis technique (LTP task). 103
Table 24. Percentage variance accounted for (PVAF) descriptive statistics and statistical comparisons for the COIs of the trial-by-trial and conditional-average ICA training approaches 124
Table 25. IDAEP task-specific outcome statistics and effects sizes for the trial-by-trial and conditional-average data 125
Table 26. LTP task-specific outcome statistics and effects sizes for the trial-by-trial and conditional-average data 126
Table 27. GA-PD and IAPD ExCrit.Epoch$_{min}$ median ASF statistics (IDAEP task) 127
Table 28. ICA-PD and COIPD (trial-by-trial approach) ExCrit.Epoch$_{min}$ median ASF statistics (IDAEP task). 127
Table 29. ICA-PD and COIPD (conditional average approach) ExCrit.Epoch$_{min}$ median ASF statistics (IDAEP task). 127
Table 30. GA-PD and IAPD ExCrit.BV$_{max}$ median ASF statistics (IDAEP task) 128
Table 31. ICA-PD and COIPD (trial-by-trial approach) ExCrit.BV$_{max}$ median ASF statistics (IDAEP task) 128
Table 32. ICA-PD and COIPD (conditional average approach) ExCrit.BV$_{max}$ median ASF statistics (IDAEP task). 128

Figure 1. Filter characteristics and terminology. 4
Figure 2. Preprocessing Optimization Pipeline (POP) structure 26
Figure 3. Signal Analysis Pipeline (SAP) overview 38
Figure 4. Wild-type grand average ERP waveform for Config.11 (IDAEP task) 51
Figure 5. Wild-type ASF for Config.11 using GA-PD (IDAEP task) 52
Figure 6. Wild GA-PD POP linear model effect size (r^2) for each configuration and their relation to A) Signal Variance (SV), B) Signal-to-noise ratio (SNR) and C) Baseline Variance (BV) (IDAEP task) 65
Figure 7. ASFs for ExCrit.BV$_{max}$ Config.13 (IDAEP task) 72
Figure 8. Grand average ERP waveform for ExCrit.BV$_{max}$ Config.13 (IDAEP task) 73
Figure 9. Grand average COI waveform for ExCrit.BV$_{max}$ Config.13 (IDAEP task) 74
Figure 10. Participant 1 ERP average waveform ExCrit.BV$_{max}$ Config.13 (IDAEP task) 75

Figure 11. Participant 27 ERP average waveform ExCrit.BV_{max} Config.13 (IDAEP task).. 76
Figure 12. Grand average ERP waveforms demonstrating the auditory evoked potentials for four different stimulus intensities in a group of strong responders. 78
Figure 13. Participant 1 COI back-projected average waveform ExCrit.BV_{max} Config.13 (IDAEP task) 79
Figure 14. Participant 27 COI back-projected average waveform ExCrit.BV_{max} Config.13 (IDAEP task) 80
Figure 15. Grand average COI waveform for ExCrit.BV_{max} Config.13 (IDAEP task) 81
Figure 16. The five ICA components of the visual evoked response, to a checkerboard stimulus 85
Figure 17. Wild GA-PD POP mean N1b amplitude potentiation effect size (Cohen's d) for each configuration (LTP task) 90
Figure 18. Wild-type grand average ERP waveform for Config.11 (LTP task) 91
Figure 19. Wild-type Grand average ERP difference waveform (Post1-Pre2) for Config.11. (LTP task) 92
Figure 20. Grand average ERP waveforms for ExCrit.BV_{max} Config.16 (LTP task). 96
Figure 21. Grand average ERP difference waveform (Post1-Pre2) for ExCrit.BV_{max} Config.16 (LTP task) 97
Figure 22. Grand average back-projected COI waveforms for ExCrit.BV_{max} Config.7 (LTP task). 101
Figure 23. Grand average back-projected COI difference waveform (Post1-Pre2) for ExCrit.BV_{max} Config.7 (LTP task) 102
Figure 24. N1b potentiation boxplots for ExCrit.BV_{max} Config.16 (LTP task). 105
Figure 25. N1b potentiation boxplots for ExCrit.BV_{max} Config.7 (LTP task). 105
Figure 26. IDAEP task: Back-projected component of interest (COI) activity for a single participant for the conditional-average and trial-by-trial ICA training approaches 122
Figure 27. LTP task: Back-projected component of interest (COI) activity for a single participant for the conditional-average and trial-by-trial ICA training approaches 123

Eq. 1 . Equation for determining the minimum value for the forgetting factor of an ocular artefact adaptive filter. 14

Eq. 2. Equation for determining the minimum amount of data required for training an ICA algorithm. 22
Eq. 3. Equation for determining filter order ($\boldsymbol{m}$). 30
Eq. 4. Equation for estimating the SNR of conditional EEG data. 36
Eq. 5. Equation for calculating the signal variance of participant conditional EEG data .. 37
Eq. 6. Equation for calculating the baseline variance of participant conditional EEG data. 37

Abbreviations List

ERP – Event-related potential

BSS – Blind source separation

PCA – Principal component analysis

ICA – Independent component analysis

HPF – High-pass filter

LPF – Low-pass filter

VEOG – Vertical electro-oculogram

HEOG – Horizontal electro-oculogram

REOG – Radial electro-oculogram

ExCrit – Exclusion Criteria

BV_{max} – Maximum baseline variance threshold

$Epoch_{min}$ – Minimum epoch number threshold

BV – Baseline variance

SV – Signal variance

SNR – Signal-to-noise ratio

SOI – Signal of interest

COI – Component of interest

ASF – Amplitude sensitivity function

POP – Preprocessing optimisation pipeline

SAP – Signal analysis pipeline

GA-PD – Grand average informed peak detection

IAPD – Individual average peak detection

ICA-PD – Independent component analysis informed peak detection

COIPD – Component of interest peak detection

Abstract

The value of EEG as a non-invasive technique for studying the time course and frequency composition of neuronal signals is well established. However, to date there is still no gold standard methodology for its analysis. Since the introduction of the technique many methodologies for artefact removal and signal isolation have been developed but their performance is often only assessed, against other methodologies, using simulated data with known and controlled artefacts and limited variance. Furthermore, these studies often only address a single stage in the entire analysis pipeline and do not consider the affect different preprocessing techniques might have upon the effectiveness of different signal analysis methodologies.

To address this issue this thesis approaches the assessment of 4 different signal analysis methodologies using real-world-data, from two different stimulus evoked potential studies, and an EEG analysis pipeline that systematically applies and adjusts various preprocessing techniques before subsequent signal analysis. This semi-automated process can be broken down into two stages.

Firstly, multiple configurations of a Preprocessing Optimisation Pipeline (POP) were performed to address three main causes of artefactual noise (1) electrical line noise, (2) non-neuronal potentials (low frequency drifts and muscle artefacts), and (3) ocular artefacts (blinks and saccades). Within the final stages of the POP data quality was assessed for each participant and poorly preprocessed participant datasets were excluded from further analysis based upon either a novel maximum baseline variability threshold criterion or a standard minimum epoch number threshold approach.

Lastly, the data was passed onto a Signal Analysis Pipeline (SAP) which estimated the amplitude of task-specific signals of interest through one of four methodologies (1) grand average informed peak detection (GA-PD), (2) individual average peak detection (IAPD), (3) independent component analysis informed peak detection (ICA-PD) or (4) component of interest peak detection (COIPD). The effectiveness of each of the different preprocessing and signal analysis strategies were then assessed based upon observing the changes within task-specific outcome statistics.

Chapter 1. Introduction

Electroencephalography (EEG) is used as a non-invasive method for recording the electrical activity of the brain. Electrodes placed upon the scalp are used to measure the small voltage fluctuations that are generated by the ion current flows of active neurons within the brain. When large numbers of neighbouring and similarly orientated neurons are active simultaneously this can create an electrical field potential, through a process called volume conduction, which can be strong enough to propagate through the brain, dura, skull and skin and be recorded by the EEG electrodes on the scalp. More specifically these electrical field potentials best represent the postsynaptic activity of, well-aligned and spatially orientated, pyramidal cells within the cortex (Teplan, 2002; Woodman, 2010). These waves of activity are often coordinated by neuronal networks to produce neural activity that oscillates at a particular frequency. The structure and function of these neuronal networks and their associated frequency activity is still something that is actively being studied today. However, over the years specific frequency ranges have been noted to have common scalp distributions and be involved in similar biological processes. This has subsequently led to various frequencies being grouped together and their combined activity being analysed in terms of a frequency band (delta <4Hz, theta 4-8Hz, alpha 8-13Hz, beta 13-30Hz, gamma 30-80Hz)(Teplan, 2002; Jia and Kohn, 2011).

In addition to the rhythmic activity that is observed in EEG, transient voltage potentials such as peaks or spikes can also be observed within the recording and are often elicited in response to events or stimuli. The negative or positive peaks within the EEG waveform are created by synchronous neuronal excitation or inhibition respectively and are referred to as event-related potentials (ERPs) (Coenen, 1995; Luck, 2005). The components within an ERP are typically characterised by their direction/polarity and their peak latency or ordinal position in relation to the event onset (Luck, 2005). For instance, when Sutton *et al.* (1965) first demonstrated that a large positive ERP component could be elicited within the EEG as a response to an unexpected stimulus, the fact that this component occurred approximately 300ms after stimulus onset led to it being termed the *P300* component. However, as this component can also be categorised as the third positive peak in the ERP waveform it is also sometimes referred to as the *P3* component. In this regard, due to the fact that peak latencies can be variable across different participants and experimental procedures it has been argued that component

naming conventions should follow the ordinal positional system for better clarity (Luck, 2005).

The temporal characteristics of ERPs have been studied for over 75 years (Davis, 1939). During this time there have been many technological and scientific advances within the area that have greatly improved the investigatory power of this technique (Collura, 1993). In its rawest and simplest form EEG data represents the amplification of neuronally generated electrical potentials as well as any 'artefact' potentials from other physiological or environmental sources. In this regard EEG data quality is often described in terms of its signal-to-noise ratio (SNR) whereby the power of a signal-of-interest (SOI), typically calculated as the amplitude variance of the signal across a given time window, is compared to an estimate of the power of the noise present within the data. A poor SNR can therefore occur as a result of excessive levels of noise in the recording or simply due to there being only a small SOI. In either case the increased amount of variance present within the signal estimate of such data has been shown to limit the strength of the statistical outcomes that can be obtained during analysis (Turetsky *et al.*, 1988; Boutros, 2008; Hu *et al.*, 2009; Widmann and Schroger, 2012). Improving EEG data quality through the use of various artefact rejection techniques and signal isolation methodologies is therefore an essential part of EEG analysis. In this regard many techniques now exist to aid in artefact removal and signal estimation. Until recently relatively little research had been conducted into the relationship between the performance of artefact removal techniques and their impact upon the statistical outcomes of real-world event-related potential (ERP) studies. However, in the last few years this area of research has started to receive a lot more attention (Hagenmuller *et al.*, 2011; Pham *et al.*, 2011; Vanrullen, 2011; Acunzo *et al.*, 2012; Cassani *et al.*, 2014; Bigdely-Shamlo *et al.*, 2015; Tanner *et al.*, 2015; Widmann *et al.*, 2015). This thesis therefore aims to further explore how artefact removal techniques may be best applied and combined within a semi-automated preprocessing analysis pipeline to achieve optimal SNR levels for real-world event-related EEG data. The effectiveness of four signal analysis methodologies will then be compared through their differences in task-specific outcome statistics as well as their sensitivity to the varying data quality levels of the different preprocessing pipeline configurations.

To enable selection of the best and most appropriate preprocessing and signal analysis techniques for inclusion within the analysis pipeline of this thesis, an in-depth review of

the current available and recommended techniques in use within EEG research was first performed and is presented here.

1.1. Preprocessing

EEG analysis is often thought of as a process that is made up of two separate stages. Broadly speaking a preprocessing stage is first used to help improve the signal quality of the EEG recording by cleaning the data as much as possible of all non-neuronal signals (i.e. noise/artefacts). The second signal processing stage then takes this data and tries to isolate the ERPs from the background neuronal signals to improve the estimation and measurement of the signals of interest. The preprocessing stage is primarily used to address three main causes of artefactual noise (1) electrical line noise, (2) non-neuronal biological potentials (low frequency drifts and muscle artefacts), and (3) ocular artefacts (blinks and saccades). Various techniques have been developed to help remove these artefacts from the recorded EEG data, with some techniques such as filtering being used to improve the quality of the data as it is being recorded.

1.1.1. *Filtering*

Within the general field of signal processing the term 'filtering' typically refers to the process of emphasising or attenuating specific frequencies within a discrete or continuous time series. In the case of EEG research this can be achieved at the point of recording, with the use of analogue (electronic) filters that are built directly into the EEG amplification circuitry or analog-to-digital converter circuitry of the recording device, or offline with mathematically constructed digital filters. The benefits of digital filters have long been understood within the engineering community, however, it wasn't until around the early 90's that the benefits of different filtering approaches started to be discussed and applied more widely within the neurophysiological community (Cook and Miller, 1992). However, the process of using filters to selectively attenuate spectral frequencies that are not associated with the SOI must also be applied with some caution as significant event-related signal distortions can also occur without careful consideration of filter design (Duncan-Johnson and Donchin, 1979; Vanrullen, 2011; Acunzo *et al.*, 2012; Chang *et al.*, 2012; Hajcak *et al.*, 2012; Rousselet, 2012; Widmann and Schroger, 2012; Tanner *et al.*, 2015; Widmann *et al.*, 2015).

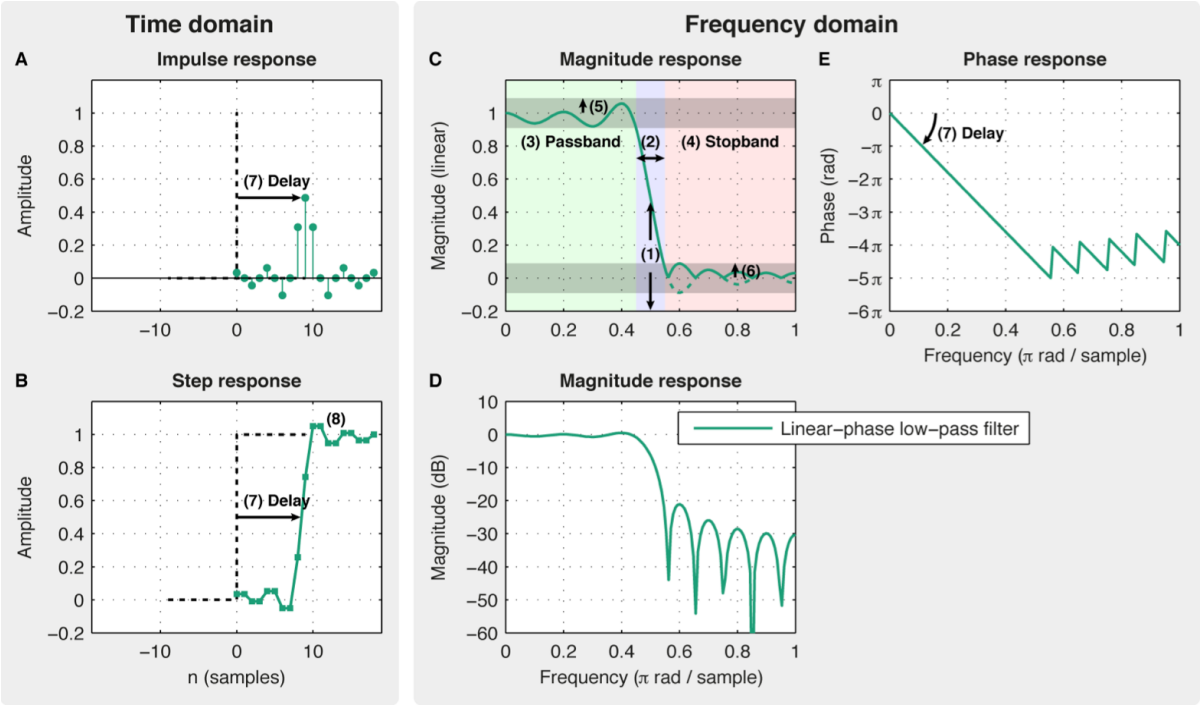

Figure 1. Filter characteristics and terminology. Time domain (A and B) and Frequency domain (C-E) responses for an example filter. Demonstrating the cutoff frequency (1), transition band (2), passband (3), stopband (4), passband ripple (5), stopband attenuation (6), filter delay (7) and ringing artefacts (8) of a linear-phase finite impulse response (FIR) filter. Copied from Widmann et al (2015)

To be able to fully understand and appreciate the positive and negative effects filtering can have upon event-related data components it is first helpful to have a rudimentary understanding of filter design. Broadly speaking there are four filter types: Low-pass, high-pass, band-pass and band stop. However, technically speaking all filters first start out as a low pass filter whereby high frequency components are attenuated within the data by a moving average. Adjusting the size of the averaging window (filter width) and the weighting of the windowed data (filter weights) enables the creation of filters with specific frequency cutoffs and attenuation characteristics. All of the other filter types are then effectively created through methods of data subtraction (i.e. a high pass filter is created by first attenuating frequencies above a given frequency cutoff value and then subtracting the resulting data from the original data to give an output where the lower frequency bands are attenuated). However, with any filter the two terms that best describe the filters' characteristics are its cutoff frequency and its transition bandwidth (Figure 1). The transition band of a given filter is often described through terms referring to its roll-off or steepness both of which are functions of the filters order (calculated as the number of filter coefficients divided by the filters length minus one).

Steep filters, i.e. those with small transition bandwidths / large orders, can be used for attenuating frequencies that closely neighbour frequencies of interest. However, this functionality comes at the cost of introducing stronger signal distortions such as temporal smearing and ringing artefacts into the data (Widmann *et al.*, 2015). In this regard, within a recent in-depth review paper of filter designs for electrophysiological data, Andreas Widmann and colleagues have devised a set of general recommendations for filter parameter selection that serve as a very good starting point for any researchers wishing to use filters as a preprocessing step for event-related EEG data.

The first recommendation was that, due to their weaker signal distortions, shallow filters (i.e. those with large transition bandwidths / small orders) should always be considered first and used wherever possible. Steeper filters should be reserved for situations where an abrupt attenuation of neighbouring frequencies is explicitly required; such as when a high-pass filter (HPF) is used to remove direct current (DC) drift artefacts due to changes in perspiration and skin conductance but where neighbouring low frequency components are crucial to the event-related potentials of interest, as would be the case for analysis of the N400 and P600 late ERPs (Tanner *et al.*, 2015).

A second recommendation that was made was that when exploratory peak analysis is to be performed (i.e. the ERPs of interest are unknown before analysis) then a high-pass filter cutoff of ≤0.1Hz and a low-pass filter cutoff of ≥40Hz should be considered. This is argued to preserve the spectral information that relates to both the shallow, low-frequency dependent, peaks of any late event-related potentials and the sharp, high-frequency dependent, peaks of any early evoked potentials (Widmann *et al.*, 2015).

Thirdly, depending upon the desired event-related signal analysis approach, it was recommended that special consideration should be given to the phase characteristics of the filter and its causality (Figure 1)(Widmann *et al.*, 2015). When a filter is applied to a dataset the filter output will inevitably be subjected to a slight delay compared to that of the filter input. How this delay is managed and how it might impact upon ERP estimation and measurement accuracy depends upon the type of filter that is used. Linear phase filters delay all of the frequencies within the passband equally. Consequently this preserves the cross spectral phase relationships within the data and helps to maintain the temporal structure of complex ERP signals. Linear phase filters are therefore primarily used when peak amplitude measures are being conducted. Furthermore, the

often large delay produced by the linear phase filter can easily be corrected by shifting the filter output back in time through a simple calculation relating to the size of the filter's impulse response. The resulting non-casual zero-phase filter can then be used in conjunction with peak latency analysis. Alternatively, a zero-phase filter can also be created through 'two-pass filtering', i.e. by running a filter first in the forward direction and then again in the reverse direction. However this approach is not recommended as the resulting filter can enhance signal distortions due to increased amounts of a filtering side-effect known as passband ripple (Widmann *et al.*, 2015). Furthermore two-pass filtering introduces additional levels of complexity that can affect cross study comparison and reproducibility as different software packages implement and report the technique in very different ways. One downfall of non-causal zero-phase filters however is that the filter output is smeared both forwards and backwards in the time domain. Therefore to be able to study signal onsets the best filtering approach to use is a non-linear minimum-phase causal filter as this approach doesn't suffer from smoothing and ringing artefacts transferring back in time (Widmann *et al.*, 2015). However, it is worthwhile to note that whereas high-pass non-linear minimum-phase causal filters introduce only a small delay, low-pass variants appear to introduce larger delays regardless of its intended minimum-phase properties (Rousselet, 2012).

Digital filters can be applied in two different ways, as finite impulse response (FIR) filters or as infinite impulse response (IIR) filters. Of the two approaches FIR filters are the most flexible and commonly used. FIR filters can have either linear or non-linear phase and produce a filter output that is defined by a limited number of data inputs. The negative effects of passband and stopband ripple can also be attenuated within FIR filters through the application of a windowing function. Selecting an appropriate windowing function is then made by considering the trade-off between increased passband and stopband ripple attenuation and transition bandwidth size (Smith, 1997; Ifeachor and Jervis, 2002).

IIR filters on the other hand have non-linear phase and therefore produce an asymmetric impulse response (i.e. ringing artefacts do not transfer back in time). The 'infinite' impulse response of the filter means that its output is in part defined by all of the preceding filter input data points. As a result its stability can be impaired from rounding errors accumulating over the course of applying the filter, with steeper transition bandwidths leading to more instability. Therefore, a key recommendation that

should always be followed when using IIR filters is that they should always be performed using at least a double-precision floating point format rather than single precision to limit the amount of rounding errors. IIR filters are often compared to analogue filters due to their shared instability and similar phase shift profile. However due to IIR filters being an offline mathematical process they can benefit from a greater amount of flexibility in filter design. As a result IIRs are often thought of as representing somewhat of a hybrid between analogue and FIR filters (Cook and Miller, 1992).

In contrast the instability and inaccuracy of analogue filters are determined by the operational characteristics of the passive and active electrical components that are used to implement the filter. For instance, every analogue filter will be coloured by the accepted tolerances and operational temperatures of the capacitors and resistors that make up the filter circuitry. Therefore, even under identical filter settings two different systems or even the same system on different days may produce noticeably different filter outputs if given the same input data. Given their limited flexibility and inconsistency analogue filters should be largely avoided, with analogue high pass filters being completely avoided in favour of recording at DC and analogue low pass filters only being applied for the purpose of anti-aliasing (i.e. using a shallow low pass filter with a cutoff frequency equal to half the analog-to-digital conversion rate therefore satisfying the Nyquist rule). In relation to this, when an EEG is recorded without a high pass filter the voltage potentials will likely exist far away from the 0μV intended ERP baseline. If an offline HPF is then applied directly to the data, the large attenuation step that would be required to bring the recording back to baseline can result in a significant and prolonged disruption of the signal. Therefore, before applying the filter in this instance it is recommended that the offset voltage is first removed from the data by subtraction (Luck, 2014). Filtering distortions such as this not only occur at the start of a recording but also occur at any point of signal discontinuity, i.e. when the EEG recording is paused and then resumed, when sections of the data are excluded or when filtering is performed upon epoched data. Therefore, the simplest way to avoid edge artefacts is to ensure that all filters are performed on the continuous rather than epoched data and that EEG recordings should be performed in one continuous recording or that a sufficient amount of time (approximately 10-30 seconds) should be included following a break in recording to allow for edge artefacts to settle before experimental conditions are performed.

Although some researchers view the many signal distortions that can occur as a result of filtering as a valid reason to avoid filtering altogether (Luck, 2005; Acunzo *et al.*, 2012), a more pragmatic and optimistic approach would be to simply view the reports as a warning for ensuring appropriate filter design. Using the default settings of a given software package or trying to apply a one-size fits all approach to filter design is when filtering is most likely to negatively impact upon signal analysis outcomes. Although very broad filter settings have been suggested as a potential way of conducting exploratory analysis (Widmann *et al.*, 2015), to achieve the very best SNR levels the filter should always be tailored with careful consideration of both, the specific signal under investigation, as well as the prevalent artefacts within the data. Furthermore there has even been evidence recently to show that different filter settings have been found to benefit classification of P50 sensory gating differently in children compared to adults (Chang *et al.*, 2012). This research therefore highlights the importance and reward of always actively investigating and assessing filter design to ensure the best possible results are being achieved.

1.1.2. *50/60 Hz ambient electrical noise*

One type of artefact that is often seen in EEG data is electrical line noise. This type of artefact can affect all of the channels within the EEG recording and presents within the data as a sinusoidal waveform that relates to the frequency of the alternating current (AC) of the mains supply. The specific frequency of the mains supply varies between different countries however typically Europe and most of Asia, Africa, Australia and South America all use a 50Hz supply however in North America a 60Hz supply is the most common (McGregor, 2016). Electrical line noise is created as a result of the AC of the mains supply creating small electro-magnetically induced voltages of the same frequency (and its harmonics) within the participant, wires and electrodes that are then summed and amplified by the EEG system. As electrical noise can be generated within any electrical conductor that is in close proximity to any mains connected electrical equipment, the primary means by which to reduce this type of noise should always be to carefully assess the recording area for potential sources of noise and seek alternatives or solutions for potentially large generators of noise. In this regard, many researchers employ various methods of electrical shielding, such as Faraday cages and specially designed EEG amplifiers with built in active shielding (ANT Neuro), to limit the effects of working within an electro-magnetically noisy environment.

Once electrical noise has been limited as much as possible at the source, the secondary means by which electrical noise can be removed from EEG data is through the use of offline data preprocessing techniques. Historically the most common approach for line noise removal was to use a band-stop filter, or notch filter, centred around the 50/60Hz line noise frequency. However, in more recent years this approach has started to be thought of as overly aggressive as either a large amount of spectral information is lost either side of the line noise frequency or significant distortions are induced in the filter output when using a band-stop filter with narrow transition band widths. Therefore, unless the spectral frequencies within the gamma band (i.e. ranging from approximately 30-100Hz) are known to be critical to the event-related signals under analysis, electrical noise can simply be removed through the use of a 30Hz low-pass filter (LPF) during data preprocessing (Luck, 2005).

However, an alternative methodology for identifying and removing electrical noise, based upon multi-taper regression, has started to gain popularity among the EEGLAB community. EEGLAB is an interactive Matlab toolbox for processing continuous and event-related EEG, MEG and other electrophysiological data (Delorme and Makeig, 2004). The use of multi-taper spectral analysis to isolate electrical noise was first proposed and performed by Mitra and Pesaran in 1999 (Mitra and Pesaran, 1999). Although their publication focused primarily upon applying the technique to magnetoencephalography (MEG) data, the technique itself was noted to be easily transferrable to other functional neuroimaging datasets. By applying the multi-taper spectral analysis technique to small sliding windows of the data they were able to obtain a slow varying estimate of the line noise's amplitude and phase, and then use this to successfully reconstruct and subtract the sinusoidal artefact from the main MEG data. In 2007, Mitra and Bokil incorporated this technique into an open source analysis platform called Chronux with the intention of providing neuroscientists with a simple software platform that can be used with a wide range of multi-channel time-series data types (Bokil *et al.*, 2010). The functionality of the multi-taper regression technique within Chronux has since been extended and incorporated into the EEGLAB analysis platform through the development of the *cleanline* plugin by Tim Mullen (Mullen, 2012). Subsequently this has led to the technique being introduced to and implemented by a greater audience as its integration within EEGLAB has made it easily accessible to a strong and active community of EEG researchers.

A very good overview of how the parameters of the *cleanline* plugin are selected and implemented by the function can be found in a recent paper by Bigdely-Shamlo *et al.* (2015). Essentially, the main default settings for taper bandwidth (2Hz), window size (4sec), slide step size (1sec) and Thomson F-statistical significance threshold ($p<0.01$) have all been noted to be sufficient for successful performance of the function upon continuous data and as such is in keeping with the advice given by the developer of the plugin. However, following extensive tests it was also found that the *cleanline* function performs poorly if the data isn't first high pass filtered or detrended (i.e. removing any linear trend present within the data). This additional step helps to improve the stationarity of the signal which consequently aids multi-taper spectral estimation through an improved level of precision within the significance testing element of the function. Once the line noise has been removed researchers can then replace the lost data from the high pass filtering or detrending to allow for more flexibility with filter design. The use of a temporary 1Hz high pass FIR filter prior to using the *cleanline* function is now widely accepted by the EEGLAB community to be the best approach for using the *cleanline* function and for line noise removal in general (Delorme and Makeig, 2004; Winkler *et al.*, 2015; Miyakoshi, 2016).

1.1.3. *Ocular artefact removal*

Ocular artefacts are a very common feature within the majority of EEG recordings, particularly within anterior EEG channels that are in close proximity to the eyes. Eye movements are observed as ocular artefacts within the EEG due to the presence of an intrinsic positive voltage potential within the iris pigment epithelial cells. This anterior voltage within the eye effectively establishes a retino-corneal dipole within the eye whereby changes in the orientation and location of this dipole results in voltage potential changes that propagate out across the scalp. There are two main types of ocular artefact that are commonly discussed in relation to artefact rejection within EEG and they are blinks and saccades. Blink artefacts are characterised by a very sharp peaked waveform, with a peak amplitude ~50-100 μV at frontal electrode sites and will often clearly stand out from the background EEG. In contrast saccades present as a step voltage change of ~16 μV per degree of visual angle in the electrodes located on the same side as the direction of the saccade. It is worth noting that to be able to clearly observe the characteristic shape of a saccade artefact it is necessary to record the EEG

without a high pass filter as otherwise each saccade step voltage will be followed by a gradual transition slope back towards the baseline voltage.

One of the simplest but arguably most aggressive ways to limit the influence of ocular artefacts within EEG data is to simply exclude any section of data that is noticeably contaminated with ocular artefacts. In practice this can either be achieved by a thorough visual inspection of the data or by systematically scanning the data for excessive voltages and excluding the surrounding data. The latter automated approach is typically only performed for blink artefacts as their characteristic large voltage potentials can be easily detected within reference ocular artefact channels such as the vertical electro-oculogram (VEOG). Blink artefacts are therefore indirectly excluded from the data when VEOG voltages in excess of +/-75 μV are detected and the surrounding data or data epoch is excluded (Luck, 2005). The effectiveness of this technique therefore depends upon setting a threshold voltage that is appropriate for the majority of ocular artefacts present within the data. However, to achieve this both the size of the ocular artefacts themselves as well as the size of the DC drifts present within the ocular reference channel needs to be considered. For instance, when large baseline variations, exist within the ocular reference channel, as can occur with changes in perspiration, this variance in the baseline voltage signal can increase the frequency of missed artefacts and false alarms. For example, when defining a voltage threshold value of +/-75 μV, a blink artefact of 100μV could be missed by the artefact rejection process if the baseline potential was to fall below -25 μV, which is entirely possible. Therefore, due to the fact that the technique has limited levels of accuracy and specificity and that it operates through endorsing systematic data loss, many researchers often try to limit ocular artefacts at their source by asking participants to regulate or restrict their eye movements. For instance, participants can be asked to refrain from blinking during experimental conditions with blinking being restricted to periods in between trials or during specified blink breaks following a block of trials. Eye movements and saccades can also be limited by asking the participant to attend to a central fixation point. Although both of these techniques help to limit signal distortions due to ocular artefacts, it is also possible that the additional cognitive load required to actively regulate a typically autonomic process may itself induce unforeseen changes within the signal (Weerts and Lang, 1973; Verleger, 1991; Ochoa and Polich, 2000).

As ocular artefact correction techniques have improved, it has enabled researchers to relax the restrictions placed upon participants relating to ocular activity and record data in a more naturalistic way (Pettersson *et al.*, 2013b). For many years, up until the late 1990's ocular artefact correction/removal was typically performed via various types of regression-based analysis (Croft and Barry, 2000c). Both time-domain (Hillyard and Galambos, 1970; Verleger *et al.*, 1982; Gratton *et al.*, 1983) and frequency-domain (Whitton *et al.*, 1978; Woestenburg *et al.*, 1983) approaches utilised an EOG reference channel to estimate artefact propagation coefficients, typically via a least mean squares (LMS) approach, for each of the EEG channels within the recording. Subsequently these coefficients would then be used to subtract a proportion of the artefact reference channel from each of the EEG channels. Furthermore, techniques incorporating both VEOG and horizontal EOG (HEOG) channels within a simultaneous multiple regression approach have since been shown to be far superior to alternative forms of simple LMS regression-based EOG correction (Croft and Barry, 2000b; Croft and Barry, 2000c).

However as different types of ocular artefacts have been shown to produce slightly different artefact propagation coefficients (Corby and Kopell, 1972; Berg and Scherg, 1991), regression techniques using a limited number of ocular reference channels may not be able to account for this additional complexity and would likely struggle to accurately correct both blink and eye movement related artefacts. One explanation that was put forward to explain the differences in propagation coefficients between different ocular artefacts was that analysis of artefacts with typically larger magnitudes, i.e. blinks, may generate a truer set of propagation coefficients as they would be less susceptible to being obscured by forward propagation of neuronal signals into the reference EOG channels and the detrimental effects of coherent interference (Croft and Barry, 1998). On the other hand, the propagation coefficients calculated for smaller ocular artefacts such as saccades would be at a greater risk of being artificially inflated by the greater relative impact of forward propagation as well as any shared physiological and environmental artefacts between the reference channels and the EEG.

In response to these concerns R.J. Croft and R.J. Barry developed the 'aligned-artefact-average' (AAA) technique, which resolved to improve the calculation of a single set of propagation coefficients by assessing artefact averaged data rather than raw EOG and EEG data. This was achieved using an initial calibration stage before performing the main EEG task/recording, where blinks and specific eye movements were made on

demand by the participant to enable a set of ocular artefact-related averages to be made. The same researchers later revised the technique to include contributions from radial EOG channels (REOG; the averaged signal from above and below the eye referenced to linked mastoids) as well as the standard HEOG and VEOG channels (Croft and Barry, 2000a). The revised AAA approach (RAAA), using VEOG, HEOG and REOG channels, has since been compared to several multiple regression techniques and has been shown to deliver a slight performance advantage over these techniques (Croft *et al.*, 2005; Pham *et al.*, 2011). This improvement in the artefact correction process has been attributed to its ability to improve EOG/EEG ratios by allowing artefact averaging to reduce the impact of forward neural propagation within the reference EOG data, before the propagation coefficients are calculated for artefact correction. Furthermore, there is some evidence to suggest that RAAA may in fact perform just as well as more modern non-linear regression based ocular artefact classification and removal techniques, such as those combing blind-source separation techniques with data from EOG reference channels (Evans *et al.*, 2012). However, due to RAAA's dependence upon both a pre-experiment ocular calibration sequence and the potentially lengthy process of offline visual inspection to select and classify ocular artefact epochs for artefact averaging, many researchers may simply prefer to use alternative techniques that provide a more automated process of artefact removal.

One methodological approach that allows for automated online artefact correction when reference EOG channels are present within the data is adaptive filtering (He *et al.*, 2004). This type of approach removes ocular artefacts from the EEG channels by iteratively shaping a set of FIR filter coefficients that are applied to EEG data and influenced by the reference EOG channel inputs. The process of calculating the best filter coefficients to use within the technique is achieved by using a recursive least squares (RLS) approach to find the best solution to minimise a target function that relates to an estimate of the output signal (He *et al.*, 2004; He *et al.*, 2007). Two parameters that affect the performance of the adaptive filter are the forgetting factor λ and the filter length M. The forgetting factor ranges from 0 to 1 and relates to how previous samples are weighted when calculating and updating the filter coefficients as new samples are fed into the system. Using a $\lambda < 1$ allows for a degree of non-stationarity between the reference EOG and the measured ocular artefact (for example as can occur over time with changes in contact impedance) by placing greater importance on more recent samples. Choosing an appropriate value for the forgetting factor, therefore, depends upon the perceived

degree of stationarity within the process. One suggestion that has been put forward, by He *et al.* (2004), to calculate an estimate for the minimum value of λ for a given recording is to first determine the average interval between two blinks and then use this value to solve Eq. 1.

$$\lambda^N = 0.5$$

Eq. 1 . Equation for determining the minimum value for the forgetting factor of an ocular artefact adaptive filter. Where λ is the forgetting factor and N is the average time between two blinks measured in datapoints.

By contrast there is no specific method of estimation for determining an appropriate filter length. Instead it has been demonstrated that the benefits of increasing the filter length noticeably diminish above M=3 (He *et al.*, 2007). Therefore due to the fact that the adaptive filter is relatively insensitive to the filter length, M=3 can typically be seen as a default parameter setting within many software approaches using adaptive filtering (Gómez-Herrero, 2007; Klados *et al.*, 2009). Interestingly, when λ=M=1 the adaptive filter has been shown to effectively operate in a similar fashion to a LMS time-domain regression approach (He *et al.*, 2007). The only difference between the two approaches relates to their implementation, whereby time-domain regression is strictly an offline process whereas adaptive filtering can be ran as an online process. However, without λ<1 or M>1 the time-domain regression-like variant of the adaptive filter would not benefit from the added flexibility afforded to the technique by these parameters. This would then limit the adaptive filters ability to deal with non-stationarity within the data as well how it copes with potential artefact waveform distortions that require a non-linear correction approach. Therefore, although adaptive filtering may represent an improvement over time-domain regression it is worthy to note that the technique doesn't resolve the issues around cross-contamination from forward propagation of neuronal signals to the ocular artefact reference channels. However, if the two techniques are first preceded by a temporary 7.5 Hz LPF, as is recommended by Gasser *et al.* (1992a) to restrict the data signals to the spectral frequencies primarily observed for ocular artefacts, then the adaptive filter may be better equipped to deal with the potential negative effects of any filter distortions. One final similarity between adaptive filtering and multiple regression that should also be mentioned is that the accuracy of adaptive filtering has also been shown to be significantly improved by the addition of an REOG reference channel (Kavitha *et al.*, 2007). It is worth mentioning that this later

point therefore also helps to further validate the argument for incorporating ocular reference channels into experimental setups (Jervis *et al.*, 1999).

Cross-contamination can cause signal loss within frontal EEG channels as well as adding false signals into other channels (Jung *et al.*, 2000a). Blind source separation (BSS) techniques have been shown to help overcome these limitations by detecting and separating neuronal and artefactual signals into a set of source components by deriving a set of spatial filters (Makeig *et al.*, 1996; Jung *et al.*, 1998a; Jung *et al.*, 1998b; Jung *et al.*, 2000a; Jung *et al.*, 2000b). Many BSS techniques exist and broadly speaking they can be separated into two groups based upon the types of statistics they use to linearly decompose the EEG into source components. Approaches based upon second-order statistics, such as principal component analysis (PCA), estimate an un-mixing matrix for the recorded signals based upon the simple assumption that the underlying sources should be uncorrelated. In the case of PCA this is achieved by calculating a rotated orthogonal un-mixing matrix for the data which effectively forces the source distributions of the components to be uncorrelated. However, as ocular artefacts and frontal neuronal signals typically have overlapping spatial topographies these components are likely to be non-orthogonal and therefore PCA may struggle to accurately isolate ocular artefacts from frontally located neural sources (Jung *et al.*, 1998b; Jung *et al.*, 2000a). In contrast SOBI (second-order blind identification; (Belouchrani *et al.*, 1997) and AMUSE (Algorithm for Multiple Unknown Signals Extraction; (Tong *et al.*, 1991)) are both BSS techniques that are also based upon second-order statistics however their use of spatio-temporal decorrelation when calculating the un-mixing matrix affords these techniques the additional flexibility to better isolate ocular artefacts (Gómez-Herrero *et al.*, 2006; Romero *et al.*, 2008). As spatio-temporal decorrelation can be thought of as a less strict version of statistical independence these two techniques are sometimes referred to as types of independent component analysis (ICA) (Hyvärinen *et al.*, 2004). However, unlike typical ICA algorithms, SOBI and AMUSE have the added advantage of both being fairly insensitive to the amount of data available for training the artefact correction algorithm (Joyce *et al.*, 2004; Gómez-Herrero *et al.*, 2006; Crespo-Garcia *et al.*, 2008b; Romero *et al.*, 2008).

Typically ICA refers to BSS techniques that use higher-order statistics (HOS) to determine statistically independent sources and operate through either maximizing non-Gaussianity, minimising mutual information or estimating maximum likelihood

(Romero *et al.*, 2008). Many different ICA algorithms exist with each one offering a slightly different approach for determining the underlying sources (Albera *et al.*, 2012). These different approaches have been designed to favour specific source distribution shapes (sub-Gaussian, super-Gaussian, symmetrical, asymmetric and multimodal) with different techniques offering better computational efficiency and reduced run times. Many comparative studies have been performed to try and identify which approach and algorithm might offer the best solution for isolating and correcting the various types of artefacts that can affect EEG recordings (Wallstrom *et al.*, 2004; Delorme *et al.*, 2007; Fitzgibbon *et al.*, 2007; Romo-Vazquez *et al.*, 2007; Crespo-Garcia *et al.*, 2008a; Albera *et al.*, 2012). However, rather than helping to establish a gold-standard approach for artefact correction the conclusions of these comparative studies are often contradictory, inconclusive or somewhat limited in their wider applicability by the characteristics of the EEG data that they choose to analyse. For instance, many studies use simulated data for testing the effectiveness of the artefact removal techniques as it allows them to directly compare the isolated artefacts against the original artefact signals used to generate the data. However, this means that the validity of the conclusions drawn from these studies is highly dependent upon how closely the simulated data actually represents the complex spatio-temporal characteristics of the true, real-world neuronal and artefactual signals. For instance, comparative studies assessing BSS techniques based upon higher-order statistics are inherently more sensitive to modelling inaccuracies due to their functional necessity for complex statistical relationships to be present within the data. Therefore, if data characteristics have been simulated poorly this can lead to an underestimation of the effectiveness of these particular techniques.

Equally, comparative studies utilising real-world artefact contaminated data are also limited in how they assess artefact correction performance. As the underlying artefact free neuronal signals of real-world data are inherently unknown they therefore aren't available for statistical comparison with the results of the technique under investigation. In an attempt to solve this issue Klados and Bamidis (2016) have recently designed a semi-simulated EEG/EOG dataset to aid in the comparison of ocular artefact rejection techniques. By combining real-world data recordings of artefact free, eyes-closed, continuous EEG with simulated ocular artefact activity, comparative studies or new techniques will be able to objectively assess performance using pre-contamination EEG signals. However, as for the assessment methodologies based upon fully simulated data, the conclusions that can be drawn from this semi-simulated data approach will also be

limited by any inaccuracies that are present in the ocular artefact model. For instance, in the model used by Klados and Bamidis ocular artefact contamination coefficients were calculated through regression analysis of VEOG and HEOG data collected in combination with eyes-open continuous EEG data from the same participants as used for collection of the artefact free data. However, by neglecting to include an REOG channel into the artefact model they may have failed to account for the potential propagation coefficient differences that likely exist for saccades and blink artefacts (Ai *et al.*; Berg and Scherg, 1991; Plöchl *et al.*, 2012). Therefore, the proposed semi-simulated EEG/EOG dataset may actually require further development of the EOG artefact model before it is widely used for assessing ocular artefact removal techniques. Furthermore, if event-related potential variants of the artefact-free data were also developed it would enable testing of each artefact correction techniques ability to handle overlapping spatial topographies of ocular artefacts and frontally generated neuronal potentials as well as artefacts that are highly correlated with event onsets.

Limitations aside, of the many previous comparative studies of ocular artefact correction techniques, those based upon SOBI BSS have been observed to consistently outperform other BSS algorithms as well as other approaches based upon regression and adaptive filtering (Joyce *et al.*, 2004; Gómez-Herrero *et al.*, 2006; Kierkels *et al.*, 2006; Romero *et al.*, 2008; Romero *et al.*, 2009; Uriguen and Garcia-Zapirain, 2015). However even though BSS techniques such as SOBI can clearly be very effective at artefact removal when no artefact reference channels are available, in their absence the burden of artefact selection and removal typically falls upon the researcher. This extra step not only increases the analytical workload but also introduces a level of subjectivity into the results that often requires a skilled professional to achieve the most accurate results. However, automated solutions for artefact component selection have also been developed which aim to achieve the same levels of accuracy through either assessing independent components (ICs) for artefact signal characteristics or by using regression or adaptive filtering techniques to correlate reference ocular artefact channels with artefact components (Wallstrom *et al.*, 2004; Gómez-Herrero *et al.*, 2006; Delorme *et al.*, 2007; Gómez-Herrero, 2007; Pettersson *et al.*, 2013a). Currently no one technique has been shown to be more accurate than the other and some researchers believe that the chance of false-positives and/or missed artefacts may be too high for serious consideration (Klados *et al.*, 2010; Klados *et al.*, 2011).

One limitation that should be mentioned with regards to the outcome of ocular artefact rejection techniques that solely rely upon BSS for artefact isolation is that even correctly identified artefact components can still have a substantial amount of neural activity present within them (Castellanos and Makarov, 2006b). Rather than being due to any form of forward propagation, as is the case in regression-based artefact removal analysis, any neural activity that is present within artefact components occurs when a suboptimal application of the technique assigns residual signal variance across all of the ICs. BSS-techniques that implement artefact removal by simply discarding the entire component can therefore result in a substantial loss of neural activity, which can then lead to significant distortions in the spectral amplitude, phase and coherence measures of neural signals.

Hybrid approaches such as wavelet enhanced ICA (wICA, (Castellanos and Makarov, 2006a)) and REG-ICA (Klados *et al.*, 2009) have therefore been developed to help resolve this issue. For instance, the wICA algorithm uses wavelet analysis to divide the BSS components into sub-components from which the low amplitude, broad spectral neural activity can first be recovered prior to artefact component deletion. The effectiveness of wICA depends upon setting a wavelet coefficient threshold that accurately estimates the magnitude of the neural broad spectral signal. Although it has been noted that this threshold value could possibly be tuned for better performance (Debnath and Shah, 2002), setting the threshold based upon a function of the data length has also been shown to perform adequately for both ocular and heart beat artefacts (Castellanos and Makarov, 2006b). In a comparative analysis with two other hybrid artefact rejection approaches, SAR and BSS-SOBI-CCA (statistical artefact rejection (Delorme *et al.*, 2007) and BSS-SOBI followed by canonical correlation analysis (De Clercq *et al.*, 2006; Gómez-Herrero *et al.*, 2006), wICA has been shown to produce the cleanest data with the least amount of signal distortion (Cassani *et al.*, 2014). Furthermore in the same study this approach was also shown to greatly improve the performance of an automated diagnostic tool for assessing biomarkers of Alzheimer's disease progression (Cassani *et al.*, 2014), therefore demonstrating the real-world benefits of the technique and artefact correction in general.

REG-ICA uses artefact reference channels and a stable version of the recursive least squares (sRLS) algorithm to adaptively filter the artefact signal out of the BSS artefact components (Klados *et al.*, 2009) rather than directly from the EEG channels as in

adaptive filtering. In comparison to LMS and wICA approaches for ocular artefact correction, REG-ICA (using the extended INFOMAX ICA algorithm and VEOG and HEOG ocular artefact reference channels) has been shown to significantly outperform the other two techniques with regards to its effectiveness at artefact removal whilst also being noted to produce fewer temporal and spectral distortions in the neural signals (Klados *et al.*, 2009; Klados *et al.*, 2011). However, further work is still required to investigate the impact of REG-ICA on ERP signal contributions as well as how the technique on the whole may be improved through the use of alternative BSS algorithms, such as SOBI, how including peri-ocular channels within the BSS training data may improve artefact isolation and how the adaptive filtering stage may be improved with additional ocular reference channels, such as REOG.

1.2. Signal Analysis Methods

Since the first major ERP-based publication in 1964 (Walter, 1964), a vast amount of EEG research has subsequently been conducted to identify and investigate the many parameters that elicit and modify a wide range of ERP components (Luck, 2005). These small voltage fluctuations within the EEG, that represent specific event-related neuronal processes, are often indistinguishable from the un-event-related background neuronal activity when viewed as raw, unprocessed, continuous EEG data. However, with the advent of electronic averaging machines in the late 1960s and with more modern advances in digital data storage and computing, many studies have now been conducted that take advantage of the reduced background neuronal activity and attenuated artefactual noise levels that can be achieved through trial-averaging methodologies (Collura, 1993). Averaging trials/epochs of ERP data together improves the SNR of the event-related data as a function of the number of trials/epochs within the average and will overall be determined by the levels of noise within the same frequency band as the components of interest (Picton *et al.*, 2000). For example, the large amount of background alpha (~10Hz) that occurs during an eyes-closed condition could significantly obscure detection of early evoked potentials such as the N100 and P200. Therefore, more trials would be required during averaging to be able to accurately detect these components. Whereas, comparatively fewer trials would be required to achieve the same signal clarity if only late evoked potentials such as the contingent negative variation or late positive potential were to be investigated, as these components are characterised by a much lower frequency spectrum then the prominent

frequencies within the background noise. The accuracy of ERP component measurements such as peak amplitude and peak latency are therefore dependent upon the levels of residual noise that are left in the data following data preprocessing (i.e. filtering, artefact removal and data rejection) and trial-averaging. As low SNR levels within individual conditional averages can make peak identification difficult, mean amplitude measurements are often used to ascertain a better estimate of the SOI when residual noise is a concern. If an ERP component's timescale has already been well defined, then the boundary limits for a mean amplitude window can be set a priori and mean amplitude measurements can be automatically calculated for the channels of interest. Alternatively, when the time frame of the SOI isn't well defined or its latency is likely to change under different experimental conditions then the boundaries of the mean amplitude window may be set based upon the peak measurements of a grand averaged conditional waveform.

1.2.1. *Trial averaging approaches*

Trial averaging and the assessment of ERP component activity can be performed in one of two ways. The simplest approach utilises the grand averaged conditional data for the study and requires the researcher to identify the ERP component peak latencies for the components of interest within this grand averaged dataset. Peak amplitude measurements for the different components are then calculated for each participant by measuring the mean amplitude across an ERP component time-window that is defined by the peak latency measurements from the grand averaged data. This approach is referred to within this thesis as grand average informed peak detection, GA-PD, where the hyphen in the abbreviated term signifies the use of a mean amplitude calculation for signal assessment. The main benefits of the GA-PD approach are its simplicity and its ability to quickly analyse large amounts of data through a semi-automated signal analysis approach. By contrast, individual average peak detection, IAPD, requires the researcher to identify the ERP components of interest within the conditional averages of each participant and as such is inherently more time consuming and requires a lot more researcher input. Furthermore, due to the fact that participant conditional averages will typically have a lower SNR than the grand averaged data, peak detection upon these averages is often more challenging and more subjective. However, by performing peak detection directly upon the participant ERP waveforms this approach accounts for any inter-individual differences in component latency and therefore should improve the

estimate of the SOI by accounting for this variance. Conversely, one type of variance that cannot be accounted for, by either approach, is that which is generated by the intra-individual differences that occur from trial to trial. Due to trial-averaging being a general prerequisite for meaningful peak detection within scalp recorded ERP waveforms, trial-by-trial analysis simply isn't possible in any significantly meaningful way with this type of data. However, BSS techniques such as those used for artefact classification and removal have been shown to be able to isolate the trial-by-trial signal well enough to enable meaningful assessment of cortical responses to singular events (Onton *et al.*, 2006; Makeig and Onton, 2011).

1.2.2. *Blind-source separation approaches*

In contrast to the well-established methodologies for analysing scalp-channel ERP averages, protocols for analysing event-related source components are constantly being created and updated. The ability of BSS to decompose the ERP signal into spatially distinct subcomponents was first demonstrated by Makeig *et al.* (1996). Under the assumption that the signals recorded at the scalp represent a linear mixture of independent and stationary source components, the INFOMAX algorithm of Bell and Sejnowski (1995) was used by Makeig's team to decompose the source component mixtures in response to detected and undetected auditory stimuli. The effectiveness of the original INFOMAX algorithm to be able to isolate both source signals and artefacts with skewed or symmetric non-Gaussian distributions has been established many times (Makeig *et al.*, 1999; Jung *et al.*, 2000a; Jung *et al.*, 2001; Delorme and Makeig, 2003; Makeig *et al.*, 2004b). Furthermore, with the introduction of the extended INFOMAX algorithm, the artefact isolation and correction properties of this approach have been further improved with additional source component decomposition functionality for sub-Gaussian spatial filter distributions, which are useful for characterising 50/60Hz line noise and natural gradient normalisation (Jung *et al.*, 1998a).

In a full-rank/complete decomposition the number of components that ICA recovers matches the number of EEG channels that are included within the recording data. Therefore, to obtain a clear decomposition of all the artefactual and event-related signal sources that are critical to explaining the dominate signals in the data it is necessary to have at least the same number of channels as expected dominant sources. In this regard, successful event-related signal isolation has been shown to be possible with as few as 14 EEG channels (Debener *et al.*, 2005), with both artefact sources and ERP subcomponents

being successfully isolated. However when performing BSS on low-density EEG data, such as this, the isolated components themselves are likely to contain a greater proportion of noise as the signal variance associated with all of the smaller non-isolated independent sources will be spread out amongst the available ICs (Onton *et al.*, 2006) as previously mentioned with regards to BSS-based artefact isolation techniques. High-density EEG systems are therefore commonly used with BSS to reduce the measurement error for the components of interest by accounting for these small independent sources. As theoretically there are a near infinite number of smaller independent sources that likely exist within the brain at all times, this means that in practice the number of channels that are used for analysis is actually calculated based upon practical concerns and limitations rather than trying to isolate every source. Primarily, as more channels are inputted into the ICA algorithm the demand upon the amount of data that is necessary for adequate training of the filter coefficients also increases. In relation to this, when Bell and Sejnowski (1995) designed the INFOMAX algorithm they recommended that it should be used with a minimum data size of 3 times the number of channels squared for a full rank analysis. However, current opinion suggests that higher density EEG recordings require a much larger coefficient (k) than Bell's recommended 3 (Eq. 2) (Onton *et al.*, 2006; Groppe *et al.*, 2009; Makeig and Onton, 2011). For instance, Onton *et al.* (2006) suggest that when performing ICA on 256 channels or more a $k \geq 20$ should be considered. However, given a typical sampling frequency of 512Hz, a 256 channel full-rank ICA analysis would require at least 43 minutes of data to meet this recommendation. Therefore for practical purposes full-rank ICA, using all of the channels available within high-density EEG systems, are very rarely conducted (Gramann *et al.*, 2010). Furthermore, as the relationship between higher sampling rates and ICA decomposition isn't yet fully understood, simply attempting to meet the recommended data size requirement by increasing the sampling rate of the recording rather than increasing the number of actual trials or continuous data that is recorded is not thought to be a viable way of improving the decomposition (Makeig and Onton, 2011).

$$Training\ Data\ > k.N^2$$

Eq. 2. Equation for determining the minimum amount of data required for training an ICA algorithm. Where *N* is the number of channels to be used in a full-rank ICA analysis procedure and *k* is a variable coefficient.

A further consideration, with regards to the data used for training the ICA decomposition, is the specific type of data that is used. Minimum data size recommendations aside, ICA has been shown to be able to effectively decompose EEG signals from conditional averaged data (Makeig *et al.*, 1997) as well as concatenated single-trial data (Jung *et al.*, 2001; Makeig *et al.*, 2002; Delorme and Makeig, 2003). The two different approaches represent two sides of a potential performance trade-off between using less data that is cleaner versus more data that has more background noise. Performing ICA on averaged datasets will minimize the impact of neural and artefactual processes that are not reliably time or phase locked to event onset (Makeig *et al.*, 1999). However, the averaging process will also decrease the temporal independence of the artefactual and non-event-related neural sources making any residual noise less likely to be isolated. Furthermore, performing ICA upon conditional averages may result in overlapping but temporally independent neuronal sources being overlooked. Alternatively, if the ICA algorithm is trained upon concatenated single-trial data this ensures that the trial-to-trial temporal independence of these components remains intact. For this reason and given the improved signal isolation that can be achieved with BSS techniques, decomposition of concatenated single trial data is often highly recommended for event-related signal analysis (Vigario and Oja, 2008).

An additional consideration with regards to the data to be decomposed is whether ICA is performed individually for each participant or if the data is concatenated across participants to form a single grand dataset with calculation of a single set of ICA weights. The benefit of creating a single set of ICA weights that can be applied to all participant datasets is that it avoids the somewhat complicated and timely processes of component clustering and enables a more straightforward interpretation of the results (Mehta *et al.*, 2009). On the other hand, component clustering requires the researcher to identify and select sets of similar components across participants to enable further analysis of event-related changes in neural activity. As the functional and anatomical organisation of the brain is likely very similar across participants without any known manifest focal pathology an 'all-in-one' approach is believed by some to be a viable alternative to individual decompositions and component clustering (Mehta *et al.*, 2009; Ponomarev *et al.*, 2010a). However, others believe that due to the non-stationarities that would be introduced into the data as a result of each participant having a slightly unique scalp topography for a given independent source, the 'all-in-one' approach would likely

negatively impact the ICA decomposition and therefore should be avoided (Zeman *et al.*, 2007).

Independent components can be assessed and clustered based upon dipole location, spectral power, event-related spectral perturbation (ERSP), inter-trial coherence, component cross-coherence, ERP contribution and scalp topography (Onton *et al.*, 2006). Depending upon the total number of components and participants within a study and the expected complexity of the underlying sources some researchers opt to simply visually assess and select components based upon the above measures (Viola *et al.*, 2011; Zhang *et al.*, 2011). Alternatively, a less subjective approach that can be applied to larger more complicated datasets utilises automated clustering strategies that can be designed with a multiple-measure approach. A good example of this approach can be seen in a recent publication by Gramann *et al.* (2010). Using the AMICA BSS algorithm (Palmer, 2006; Palmer *et al.*, 2006; Palmer *et al.*, 2007) to decompose EEG data from a visual oddball task the researchers were able to systematically isolate approximately seven components for each participant that related to the P300 event-related potential. EEG data from a high-density (248 channel) recording system was first sorted based upon measures of channel quality which resulted in approximately 130 channels on average remaining for each participant and subsequent BSS. ICs with dipole source locations outside of a spherical forward head model were categorised as sources of non-neural artefact and excluded from further analysis. Then multiple measures of event-related dynamics were calculated for the remaining ICs and from these 10 principal component measures were selected by PCA for subsequent clustering. Finally, component measures relating to dipole location and ERSP were up-weighted prior to clustering to maximize its sensitivity to IC ERP differences and visual-task-specific classifiers. This highly tailored approach to clustering resulted in the high-density EEG data being reduced to approximately 7 task related neuronal sources for each participant. The activations of these ICs were then back-projected on to scalp channels where standard ERP peak analysis was performed upon the summed waveform of the visual evoked response.

Due to the general uniqueness of each study dataset, with regards to the specific signals being analysed and the varying levels and contributions of different types of artefactual noise present within the data, it is highly likely that there is no one best strategy for clustering and that each new situation should be met with careful consideration of the

important factors to be assessed within the experimental protocol. Component clustering methodologies are now offered within EEGLAB, however, only a minimal amount of advice currently exists within the EEGLAB community for how these clustering approaches might be best applied. Furthermore, the use of BSS techniques and clustering approaches for the study of event-related potentials is far too often used exclusively as a method of 'exploring' the dynamics of the ERP signals rather than for directly performing statistical tests upon the signal components themselves. Therefore, to my knowledge there are currently no studies that have directly compared how BSS techniques may be used to improve the assessment of ERP signals and the statistical outcomes of ERP tasks, with simple trial-averaging approaches.

The benefits of performing various preprocessing techniques upon raw datasets prior to signal analysis are widely promoted in the EEG community (Picton *et al.*, 2000). However, the degree to which these techniques are able to improve the SNR when combined together and their relative importance towards improving signal estimation via various signal analysis methodologies has thus far received little consideration. Furthermore, when a new artefact removal technique or signal estimation approach is first developed its capabilities are typically assessed using simulated datas ets. The aim of this work was to primarily assess the effectiveness of 'complex' ICA-based signal analysis approaches in comparison to 'simpler' event-related trial averaging ERP approaches, using real world data. A secondary aim was then to investigate the sensitivity of each signal analysis approach to the different data preprocessing strategies. These aims were investigated using data from two different passive sensory evoked potential tasks and data analysis was performed using custom analysis scripts that iteratively and systemically applied different preprocessing and signal analysis strategies to the data.

In the following chapter the general organisation and design of these custom scripts will be explained in relation to a set of preprocessing and signal analysis pipelines that can be configured to offer different analysis strategies. This will then be followed by two separate chapters for each of the passive sensory event-related potential tasks, where the history and methodology of each task will be explained before the task-specific signal analysis and statistics are reported and discussed in brief. The final chapter will then review the findings from both tasks, whilst discussing any limitations and areas for potential future work.

Chapter 2. Analysis pipeline configurations

Data preprocessing within this thesis was conducted utilising a set of artefact removal techniques organised into a preprocessing optimisation pipeline (POP) (Figure 2). Signal and noise level estimates were then calculated for each of the outputs from the many different POP configurations before task-specific signal analysis was performed on each of the preprocessed datasets. Task-specific event-related signal analysis was performed on the preprocessed datasets by four different signal estimation/isolation approaches, which are referred to collectively as the signal analysis pipeline (SAP). The impact the different POP and SAP configurations have upon the task-specific outcomes were then assessed through observing and interpreting changes in statistical significance and effect size of task-specific outcomes.

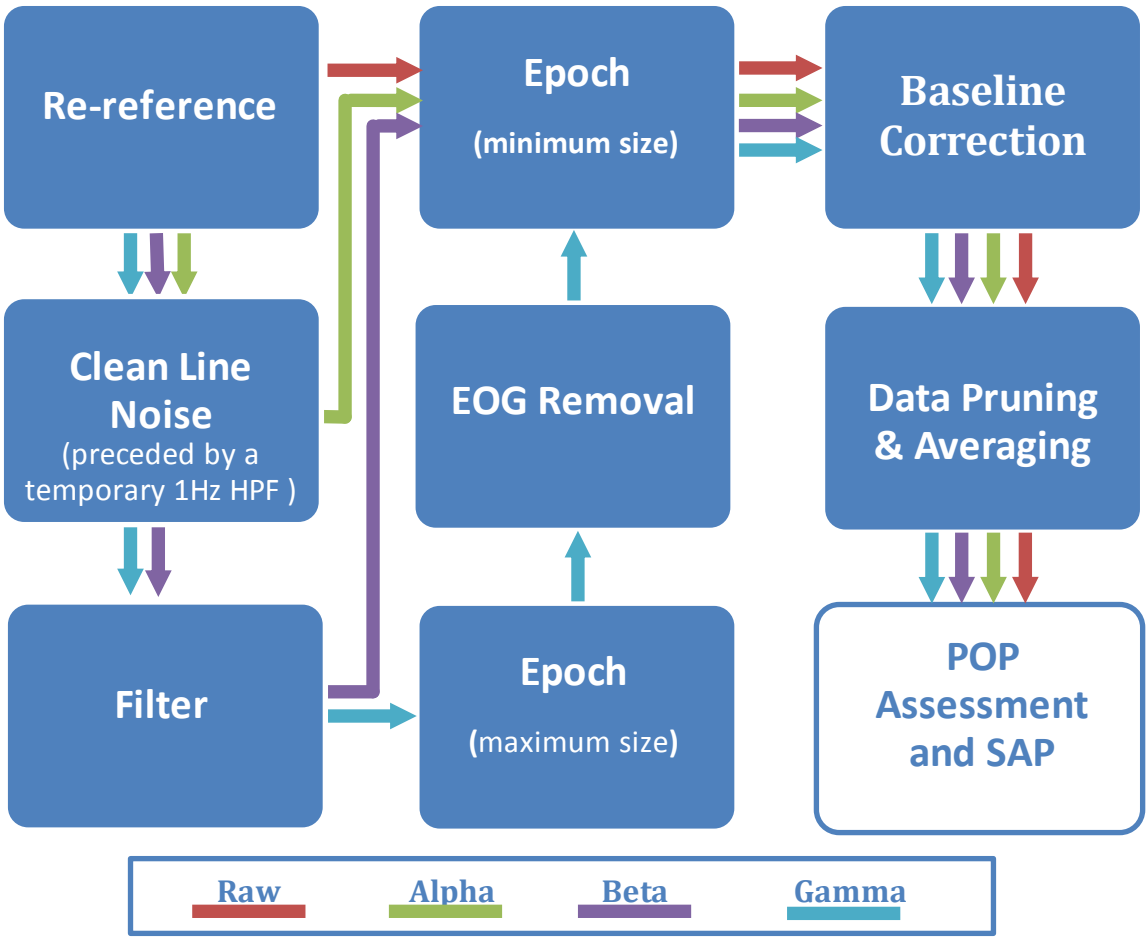

Figure 2. Preprocessing Optimization Pipeline (POP) structure. Where the workflow through the different preprocessing stages for the different POP configuration types (i.e. raw, alpha, beta and gamma) is illustrated by the different coloured arrows. The final stage in the pipeline represents the point at which the data is ready for the data quality assessment and the signal analysis pipeline (SAP) stages.

2.1. Preprocessing optimization pipeline (POP) workflow

The POP utilises different sets of artefact removal functions and parameters to systematically determine the best preprocessing strategy for each study. For ease of interpretation the main differences between the different configurations are referred to by four POP configuration types: raw, alpha, beta and gamma (Table 1). As the optimization process is performed, the raw data is passed through the different POP configurations multiple times, each time following the specific POP pathway defined by its configuration type in combination with the specific preprocessing parameter settings for that configuration number (Figure 2).

Table 1. POP configuration soft parameter settings. Data preprocessing was typically conducted in 16 different ways. With each POP configuration type (Raw, Alpha, Beta, Gamma) including elements from the prior preprocessing stages. Note when no filter cut-off value is present in the table this refers to when no digital filtering was applied and instead the amplifier filter settings at the time of recording applies to the data i.e. LFP 0.1Hz and HPF 100Hz. The gamma type POP configurations include two different types of ocular artefact removal (OAR) based upon either least mean squares (LMS) regression or conventional recursive least squares (CRLS) regression.

Configuration Number	Configuration Type	Clean Line Noise	HPF Cutoff (Hz)	LPF Cutoff (Hz)	OAR Type
01	Raw	×	×	×	×
02	Alpha	✓	×	×	×
03	Beta	✓	0.5	×	×
04	Beta	✓	1.0	×	×
05	Beta	✓	×	30	×
06	Beta	✓	×	60	×
07	Beta	✓	0.5	30	×
08	Beta	✓	0.5	60	×
09	Beta	✓	1.0	30	×
10	Beta	✓	1.0	60	×
11	Gamma	✓	1.0	30	LMS
12	Gamma	✓	1.0	30	CRLS
13	Gamma	✓	0.5	30	LMS
14	Gamma	✓	0.5	60	CRLS
15	Gamma	✓	×	30	LMS
16	Gamma	✓	×	30	CRLS

2.2. Preprocessing techniques and parameter selection

Data preprocessing within this thesis, focused upon three artefact removal techniques that could all be implemented within an automated EEG analysis pipeline. All of the preprocessing and signal analysis pipeline configurations were developed using Matlab R2014a (The MathWorks, Inc.) with additional code and functionality from the EEGLAB toolbox (Delorme and Makeig, 2004) and its associated plugins. The different preprocessing techniques incorporated into the POP addressed three main causes of artefactual noise: (1) electrical line noise, (2) non-neuronal biological potentials (perspiration and muscle artefacts), and (3) ocular artefacts (blinks and saccades). After reviewing a wide range of possible techniques, as detailed in the introductory chapter, those shown to be the most effective at their respective artefact removal role where chosen for inclusion within the POP. Specifically the three artefact removal techniques incorporated into the POP utilised: the *Cleanline* plugin developed by Tim Mullen (2012), a range of high and low pass filters as recommended by Widmann *et al.* (2015), and ocular artefact removal functions from the Automatic Artefact Removal (AAR) toolbox plugin by German Gómez-Herrero (2007). It is necessary to note, however, that the techniques selected from the AAR plugin were not the first choice for ocular artefact removal. As detailed in the introductory chapter hybrid ICA-based ocular artefact removal techniques, such as REG-ICA, have been shown to outperform simple regression techniques ((Klados *et al.*, 2009; Klados *et al.*, 2011)). Furthermore it has been argued that these techniques will work best if they incorporate either the SOBI or AMICA ICA algorithms for signal decomposition (Uriguen and Garcia-Zapirain, 2015). However, when trying to incorporate the REG-ICA plugin or SOBI BSS from the AAR plugin into the analysis pipeline, an unknown error in both of the functions was found to cause these processes to fail. After many attempts to identify and resolve the problem, the decision was made to utilize the simple regression analysis techniques within the AAR plugin for ocular artefact removal within the POP.

2.2.1. *Cleanline plugin*

The POP was setup with a combination of hard and soft parameters. A soft parameter relates to any value or processes within the POP that may be changed to potentially improve the effectiveness of the artefact removal process that the parameter targets. For example, the different types of ICA algorithm that can be used within various artefact removal techniques can be thought of as a soft parameter. Whereas, a hard parameter refers to any process or value which is believed to be appropriate for the majority of studies and can typically be thought of as 'best practice'. Hard parameters are therefore set values within the different POP configurations. For instance, the *cleanline* plugin follows a set of hard parameters whereby the default settings present within the function are expected to be sufficient for the majority of cases (Mullen, 2012). The effectiveness of the default settings (i.e. taper bandwidth (2Hz), window size (4sec), slide step size (1sec) and Thomson F-statistical significance threshold ($p<0.01$)) and the function as a whole has only recently been improved upon through the notable addition of a 1 Hz HPF before running the *cleanline* function (Bigdely-Shamlo *et al.*, 2015). However, to ensure that researchers aren't restricted in their future choice of filtering strategies this 1Hz filter is only used temporarily to improve the stationarity of the line noise signal before its estimation and elimination. Essentially this is achieved by capturing the line noise signal that the *cleanline* function removes and subtracting it from a copy of the EEG data taken before the 1 Hz HPF was applied.

2.2.2. *Filtering*

When considering the main filtering parameters to be incorporated into the POP, the impact filtering can have upon the underlying signal should always be considered. The detrimental effects of inappropriate filtering have been widely documented (Vanrullen, 2011; Acunzo *et al.*, 2012; Chang *et al.*, 2012; Widmann and Schroger, 2012; Tanner *et al.*, 2015; Widmann *et al.*, 2015) and discussed within the introductory chapter of this thesis. Therefore, to ensure that an adequate filtering solution was applied to the data an exploratory set of filters was incorporated into the POP and performed upon the continuous EEG. The impact of different filter cutoff combinations for the high pass and the low pass filters was then assessed through changes in signal and noise related data outcome measures. As previously discussed, this approach is highly recommend by Andreas Widmann's team in their recent paper (Widmann *et al.*, 2015) which systematically demonstrated the pitfalls of inappropriate filter selection and

recommended strategies and best practices for successful digital filter design. For the purpose of this thesis a FIR filter design with a Hanning window was chosen based upon these recommendations. Filter-induced signal distortions that could affect event-related peak amplitude detection were limited through the use of zero-phase FIR filters with steep high pass transition bands and shallow low pass transition bands. To achieve the desired steepness of each filter the transition bandwidths were set to 0.1 Hz and 10 Hz for the high pass and low pass filters respectively. The order of each filter was then calculated as per equation (Eq. 3).

$$m = \Delta F / (\Delta f / f_s)$$

Eq. 3. Equation for determining filter order (m). Where ΔF is the transition width for a given window type (3.1 for a Hanning window), and f_s is the sampling frequency.

When designing the exploratory filter set to be used in the POP the filter orders were kept constant for both the high and low pass filters (HPF: $m = 15500$, $\Delta f = 0.1\text{Hz}$, LPF: $m = 156$, $\Delta f = 10\text{Hz}$). However, the filter cut-off frequencies were altered for each filter to create 8 different filter combinations of high and low pass filters (Table 1). The impact of the 8 different filtering strategies upon the effectiveness of data processing and signal analysis outcomes were assessed as a part of the beta-type POP configuration. The filtering strategies that give rise to the best outcomes were then applied to the next stage in the preprocessing optimization process. Whereby, the effects of different ocular artefact removal (OAR) routines were tested upon datasets that had already undergone data preprocessing using the *cleanline* function and a selection of the best filtering strategies.

2.2.3. *Ocular artefact removal*

As previously described many ocular artefact removal techniques have been designed and tested upon simulated and real world EEG data and of these the hybrid ICA-based approaches appear to offer the most ocular artefact removal potential. However, as the hybrid ICA-based approaches within the REG-ICA (Klados *et al.*, 2011) and AAR (Gómez-Herrero, 2007) plugins could not be successfully incorporated into the POP design, two simple regression-based adaptive filtering techniques within the AAR plugin were used instead. The Least Mean Squares (LMS) and Concurrent Recursive Least Squares (CRLS) regression algorithm based approaches were specifically selected for the fact that they are known to demonstrate opposing strengths and weaknesses. To be more specific the LMS algorithm is arithmetically simple and is considered to be numerically stable when it is used with a small enough learning step, however this also means that it is often slow to converge upon the optimal adaptive filter weights. Conversely, the CRLS algorithm can better cope with non-stationary ocular artefact data due to the inclusion of a forgetting factor within the algorithm, whilst being faster to converge upon the optimal filter weights. However this approach can cause an accumulation of round-off errors in the filter weight approximations leading to numerical instability (Gómez-Herrero, 2007).

Both techniques used the HEOG and VEOG channels for artefact referencing and each technique was implemented using all of the default parameter settings present within the AAR plugin with the exception of the learning rate (mu) being decreased for the LMS function from 1e-6 to 1e-8 as some participant datasets failed to reach convergence with the higher learning rate. Keeping the filter order of the regression techniques set at the default level(m=3) is supported by the fact that even though increasing the filter order would allow for more ocular artefacts to be removed it would also increase the risk of over correction (Gómez-Herrero *et al.*, 2006). Furthermore, it has been demonstrated that the benefits of increasing the filter order, as measured by improvements in the mean squared residual error of artefact corrected EEG, noticeably diminish with filter orders above m=3 (He *et al.*, 2004).

Rather than performing the techniques upon the continuous EEG data, which would require the manual exclusion of any section of the data that demonstrated uncharacteristically high levels of noise (as often occurs at the beginning and end of a task and during any rest breaks, when ocular and muscular artefacts are exaggerated),

the techniques were instead performed on the concatenated task-specific epoched data. However, to ensure the maximal amount of task data was still available for the OAR techniques to work with, conditional epochs were created that were as large as possible whilst ensuring that the data from neighbouring trials did not overlap. For instance to achieve this for a simple auditory oddball task where a variable ISI of 1500ms +/- 100ms was used, then the maximum epoch size for subsequent OAR would be created by epoching around each stimulus from -200ms to 1200ms (with the smallest possible ISI being 1400ms). The specific details relating to the time frames used for defining the maximum epochs within this thesis are task-specific and will therefore be detailed in the methodology sections of each task later in this thesis. Following performance of the ocular artefact removal function all of the epochs from both tasks were trimmed down (i.e. the minimum epoch size of -200ms to 500ms was used) to shift the focus to the time frame associated with the expected early evoked potentials for the remaining preprocessing and signal analysis stages.

2.2.4. *Baseline correction*

As peak amplitude measurements of ERP activity are typically performed by measuring the voltage of the ERP in comparison to the baseline activity, it is necessary to first baseline correct each epoch to enable valid peak detection measurements and comparison both within and across participant datasets. For stimulus-locked epochs the recommended approach for baseline correcting data is to subtract the average voltage of the 200ms preceding stimulus onset from all of the data points within the epoch (Luck, 2005). Performing baseline correction for each epoch based upon the pre-stimulus interval ensured that all of the epochs shared a common baseline that was free from being influence by any subsequent stimulus related activity or DC drift. However, it is worth noting that the pre-stimulus interval can be biased by the ERP activity of preceding stimuli when the interval between subsequent stimuli is very small (i.e. < 1sec) (Woldorff, 1993). Furthermore, neuronal activity associated with anticipatory processes can also affect the pre-stimulus interval (Luck, 2005; Woodman, 2010). However, given that the shortest interval between consecutive stimuli in either of the tasks within this thesis is greater than 1 second and that both tasks are passive in nature, the 200ms pre-stimulus interval was believed to be free from any potentially baseline biasing effects and was therefore used for all baseline correction procedures within the POP.

2.2.5. *Data pruning by epoch exclusion*

As the different preprocessing strategies are not guaranteed to correct all of the artefacts that are present within the data, data pruning is therefore often used to ensure that any data with large amounts of noise/artefacts are excluded before the data is sent on to the signal analysis stage. One methodology by which this can be achieved assesses all of the epochs for each participant and any epochs that are found to contain excessively large voltages are excluded from further analysis (Hoehl and Wahl, 2012). As neuronal activity typically doesn't exceed +/- 50 μV (Aurlien *et al.*, 2004) any epochs that contain voltages that exceed this range are likely to still contain large artefacts following unsuccessful artefact correction. The first stage of data pruning within the POP therefore utilised an epoch exclusion approach based upon identifying epochs with voltages exceeding +/- 50 μV in relation to the baseline.

2.2.6. *Averaging*

Following the first stage of data pruning all of the remaining epochs for each participant were averaged together to make task-specific conditional averages for each participant. The process of creating conditional averages can essentially be considered as another method of artefact correction. This concept is based upon the premise that if the activity of the event-related signal acts as a constant within the event-locked epoch, then if non-event-related neuronal activity is simply thought of as a random process with varying phase, then as multiple epochs are averaged together the signal will remain intact whereas the phase differences in the activity relating to the neuronal noise will result its own cancellation (Picton *et al.*, 1995). The mathematics behind the noise cancellation properties of the averaging process can be explained by the squared-root law. This law describes how the attenuation of noise is directly proportional to the square root of the total number of epochs (Callaway, 2012). However, although the total number of epochs available for averaging is understood to be a very important factor in improving the SNR of EEG study datasets, the threshold for the minimum number of epochs that is believed to represent a sufficiently large enough data sample to ensure data quality varies broadly across the research community. For example, many EEG researchers attempt to ensure the quality of their data by setting participant/data exclusion criteria that are based upon requiring a minimum number of epochs to be present within the conditional averages (Csibra *et al.*, 2001; Woodman and Luck, 2003; Lin *et al.*, 2013; Nikolaev *et al.*, 2013; Kaatiala *et al.*, 2014). However, this methodology is potentially flawed as it does

not directly assess the SNR of the conditional averages or any other measure relating to the cleanliness of the data. The necessity for there to be a more direct way of assessing and ensuring signal quality during EEG analysis and preprocessing was recognised many years ago in a set of EEG analysis guidelines published by Picton *et al.* (2000). In this paper Picton and his team suggest that experimenters should endeavour to produce data whereby the level of the noise is known to be lower than the expected signal differences of interest. However, no specific methodology was put forward for how this could be assessed or achieved. Therefore, in this thesis participant exclusion will be undertaken through both a standard minimum epoch threshold approach as well as through the introduction of a novel maximum baseline variability (BV) threshold approach that is more in line with Picton *et al's* recommendations.

2.2.7. *Data pruning by participant exclusion*

Within the second stage of data pruning all of the datasets from the different POP configurations underwent two different versions of participant exclusion. The standard minimum epoch threshold approach (ExCrit.$Epoch_{min}$) required each participant to possess at least 20 epochs in every test condition to remain within the POP configuration study dataset. Whereas, the novel maximum BV threshold approach (ExCrit.BV_{max}) first estimated the size of the task-specific event-related signal change and used this to define a baseline variance exclusion threshold measure. This was then subsequently used to directly assess the quality of each of the conditional averages for each participant and therefore ensured that any conditionally related change in the SOI was not likely to be masked by high levels of noise.

To determine a reasonable estimate for the task-specific SOI the raw EEG data for each task was first preprocessed using a wild-type variant of the POP whereby all epochs and participants were kept for subsequent analysis. This unique preprocessing approach ensured that data from all participants was being considered when determining the ExCrit.BV_{max} threshold value and that the benefits of signal averaging and the square-root law were being applied equally across all participants. The mean SNR levels for each wild-type POP configuration dataset were then assessed (Eq. 4) and the gamma-type POP configuration that produced the largest SNR was selected for subsequent analysis.

The amplitude of the SOI for each task was estimated by assessing the grand averaged conditional ERP waveforms for the selected POP configuration dataset. A GA-PD approach was used based upon its practicality and simplicity when handling datasets with the potential for excessive levels of noise. Under this methodology manual peak detection was performed on the conditional grand averaged data for each task, with the peak latencies being noted for each of the relevant ERP components (i.e. the peak latencies of the N1 and P2 auditory and visual evoked potentials for each task). The conditional grand-averaged peak latencies were then used to define mean amplitude time windows for the task-specific event-related SOIs (e.g. the N1-P2 mean amplitude difference for the auditory task) and a measure of the SOI for each participant's conditional averaged data was calculated.

A measure of the size of the task-specific event-related signal change (e.g. the change in N1-P2 mean amplitude difference between two visual task conditions) was calculated for each participant using the data from the GA-PD technique (the task-specific details of this process are further described in the methodology sections of the individual task chapters) and the median value across participants was taken as a robust estimate of the size of the task effect. The baseline variability exclusion threshold was then set to less than the square of the median task effect value. The BV for the standard pruned POP configuration datasets was then calculated for the conditional averages of each participant (Eq. 6) and any participant that exceeded the BV exclusion threshold for any of their conditional averages were excluded from further analysis.

2.3. POP Assessment

The POP assessment stage consisted of measuring three indexes of signal and noise strength from the datasets of each of the different processing strategies. The SNR, BV and signal variance (SV) index measures were all calculated to help understand the effectiveness of the different preprocessing strategies, as well as how they might impact upon and relate to any differences in the signal analysis outcomes. To be able to calculate the three assessment indexes the epoched conditional data for each participant from the various POP configurations and participant exclusion methodologies were each assessed. The three assessment indexes were then calculated for each participant and each condition separately with all of the subsequent conditional indexes being averaged together at the end of the process to give three single assessment indexes that represent each of the different preprocessing strategy datasets for each task. For ease of

interpretability it was at this point when a single channel of interest was selected for each task. The reporting and interpretation of the assessment indexes and any further analyses were then restricted to this single channel.

2.3.1. *Signal-to-noise ratio*

The mean SNR for each dataset was calculated using a well-established equation taken from the work of Möcks *et al.* (1984)(Eq. 4). Through this equation an estimate of the SNR ($\widehat{SNR}$) for the conditional data of each participant, POP configuration and exclusion criteria group was established by first estimating the power of the noise ($\widehat{PN}$) within the conditional data of each participant, followed by estimating the power of the signal ($\hat{PS}$) within each participant's conditional average. The $\hat{PS}$ was calculated by first measuring the amount of variance in each participant's conditional average and correcting this value by an estimate of the residual noise that was based upon the total number of epochs within the conditional average for each participant. All of the $\widehat{SNR}s$ for each participant and condition within a specific POP configuration and participant exclusion criteria group were then averaged together to give an index of the SNR for that dataset. These SNR indexes were then used as the primary index by which the different preprocessing strategies were assessed.

$$P(\bar{x}) = \frac{\sum_{t=1}^{T}(\bar{x})^2(t)}{T}$$

$$P(x_i - \bar{x}) = \frac{\sum_{t=1}^{T}(x_i - \bar{x})^2(t)}{T}$$

$$\widehat{PN} = \frac{\sum_{i=1}^{N} P(x_i - \bar{x})}{(N-1)}$$

$$\hat{PS} = P(\bar{x}) - \frac{\widehat{PN}}{N}$$

$$\widehat{SNR} = \frac{\hat{PS}}{\widehat{PN}}$$

Eq. 4. Equation for estimating the SNR of conditional EEG data. Where x is the conditional epoch voltage data for a participant, T is the signal time window (0ms to 500ms), and N is the total number of conditional epochs

2.3.2. *Signal Variance*

The SV index was calculated to give an idea of the general size of the ERP for a given task and was simply calculated as the variance of the conditional average which in effect was calculated during the first step of the SNR calculation (Eq. 5).

$$SV = P(\bar{x})$$

Eq. 5. Equation for calculating the signal variance of participant conditional EEG data

2.3.3. *Baseline Variance*

The BV for each participant's conditional data was also calculated using the same mathematical formula that was used to assess the power of the conditional average signal $P(\bar{x})$. However, the BV calculation was performed upon the time frame relating to the baseline window of the conditional epochs (i.e. -200ms to 0ms). The BV index was calculated for each preprocessing strategy dataset as the average of all participant conditional BVs within the dataset and was used to assess the general cleanliness of the data (Eq. 6).

$$BV = \frac{\sum_{t=1}^{T_{baseline}} (\bar{x})^2(t)}{T_{baseline}}$$

Eq. 6. Equation for calculating the baseline variance of participant conditional EEG data. Where $T_{baseline}$ is the baseline time window (-200ms to 0ms)

2.4. Signal analysis techniques and parameter selection

The impact that different signal analysis methodologies have upon the measurement and estimation of the event-related signal and the magnitude of the task-specific statistical results was explored for 4 different ERP and ICA based signal analysis techniques. A general description of these 4 approaches can be found in Figure 3. For all of the datasets for both of the tasks, each of the 4 approaches required the identification of the N1 and P2 peaks in response to the different auditory and visual stimulus conditions. For each of the different signal analysis approaches peak detection was performed manually for each task-specific stimulus condition. Within the auditory task the N1 and P2 peak data was used to calculate an N1-P2 peak amplitude difference measure for each condition. In the visual task, identification of the N1 and P2 peak latencies was used to define a time window for classification of the N1b (defined as the time region from the N1 peak to the halfway point between N1 and P2) before calculating its mean amplitude. The overall impact of the different preprocessing strategies and the four signal analysis approaches were then assessed in relation to how they affected the strength and significance of the task-specific statistical tests that were performed upon the event-related signal analysis output measures.

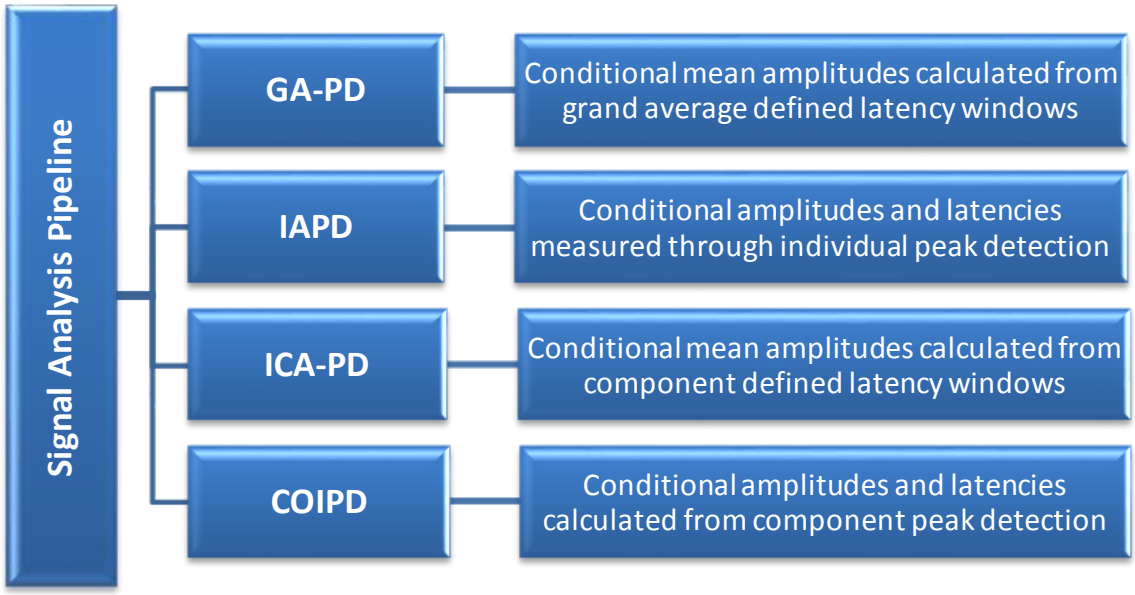

Figure 3. Signal Analysis Pipeline (SAP) overview of ERP-based (GA-PD and IAPD) and ICA-based (ICA-PD and COIPD) signal analysis approaches

2.4.1. *Grand average informed peak detection (GA-PD)*

The general process by which GA-PD is performed has already been detailed above due to its involvement in determining the ExCrit.BV_{max} threshold value. The only notable difference being that for the main GA-PD analysis was performed upon the standard pruned POP configuration datasets as opposed to the wild-type variant data as described previously. The main advantage of using this GA-PD technique comes from its simplicity. The process of performing peak detection upon the grand average enables those with less experience to perform the task with greater ease as the grand averaging process helps to further improve the signal-to-noise ratio making peak classification easier. However, this approach comes at the cost of losing data related to the inter-individual differences in peak latency. To address this, the mean amplitude of an ERP component is often calculated for each participant by averaging across a time-window centred around the grand average peak latency information.

In the case of the auditory task the N1 and P2 mean amplitudes for each individual were calculated from the conditional averages using a 20ms time window around each grand average defined peak latency. Narrow time windows such as this are often used when studying early evoked potentials such as the N1 and P2 (Amitay *et al.*, 2013; Schroder *et al.*, 2014).

In the case of the visual task the N1b mean amplitude for each individual was simply defined by the grand average peak latencies without any additional flexibility added to the time window. This approach was followed as extending the grand average defined N1b time window to allow for inter-individual latency differences would have ran the risk of shifting the focus of the analysis to any potential conditional changes in the N1 peak amplitude.

Due to the overgeneralised peak amplitude measurement approach taken by the GA-PD procedures it was hypothesised to be the least sensitive to POP configuration changes and was therefore believed to be the most likely to produce the lowest values for task-specific effect sizes and levels of significance.

2.4.2. *Individual average peak detection (IAPD)*

As the name suggest IAPD involves performing manual peak detection upon the conditional averages of each participant. Depending upon the size and complexity of the study and the signal quality of the data this can be a very lengthy process. Furthermore,

a significant downside to this technique is the inherent level of subjectivity that is present within the act of personally defining each peak for all of the different participant datasets. However, when analysis is performed by a skilled professional the issues surrounding subjectivity can be limited and the benefits of the technique to account for potential inter-individual differences in peak latencies can be achieved in its place. Consequently, the statistical outcome measures of each task may then benefit from this reduction in measurement error.

2.4.3. *ICA informed peak detection (ICA-PD)*

To be able to perform either of the ICA-based analysis techniques the EEG datasets from the different POP configurations for each study must be decomposed into a set of ICA components. Based upon the current research and guidelines described in the introductory chapter, the ICA signal decomposition for each participant dataset was trained upon a non-baseline corrected variant of the POP output using the extended INFOMAX ICA algorithm (Jung *et al.*, 1998a). Under this algorithm, ICA components can be separated based upon a wider range of Gaussian distributions. Therefore, this ICA algorithm should be able to better isolate the typical super-Gaussian source distributions indicative of neuronal signals, as well as the sub-Gaussian source distributions that are often observed for various types of artefactual noise. The effectiveness of ICA is highly dependent upon the quality and quantity of data it is given to work with. In relation to this, a gold standard is yet to be defined. For example, some researchers prefer to train the ICA decomposition using concatenated trial-by-trial epoch data for each participant (Debener *et al.*, 2005; Onton *et al.*, 2006; Gramann *et al.*, 2010; Viola *et al.*, 2011). Whereas others have chosen to train their ICA algorithms using concatenated conditional average data for each participant (Vigário *et al.*, 1998; Makeig *et al.*, 1999; Bishop *et al.*, 2011). On one side of the argument conditional average data can be argued to offer a better quality of data due to the associated high SNR levels of averaged data. However, the counter argument to this is that the averaging process itself will limit the number of data points available for the ICA algorithm to learn from. As both sides of this argument appear to be equally valid the decision was made to perform the ICA-based signal analysis using both techniques so that the effectiveness of each approach could be investigated. The results of this comparative analysis can be found within the appendix section of this thesis (Chapter 6.1, page 119). Based upon the conclusions drawn from this supplementary investigation a decision was made to focus

upon the results of the concatenated trial-by-trial epoch data approach. Furthermore, due to ICA analysis being a computationally and time intensive process the focus of analysis was further restricted to just a subset of the best raw, alpha, beta and gamma POP configuration outputs.

Once the ICA components were generated, the component that best represented the activity of the SOI was selected for further analysis. This process was achieved through the use of a custom designed automated component selection process that calculated the percentage of variance within the conditional average ERP that was accounted for by each component when its activation waveform was back-projected upon the channel of interest. Calculation of the 'percent of variance accounted for' measure (PVAF) was adapted for each study by restricting the time-window of the calculation to a time frame that was based upon the GA-PD peak latency results from the ExCrit.BV_{max} threshold value setting stage.

In the case of the auditory task, the PVAF time window was designed to represent the entire time range from the N1 peak to the P2 peak and was defined by calculating the mean of the peak latencies for the conditional averages of the wild-type GA-PD results with the addition of 20ms either side to account for inter-individual differences. Similarly, the PVAF time window for the visual task was determined by calculating the mean N1b time frame across the different conditional averages for the wild-type GA-PD results with the addition of 20ms either side. It is worth noting that whereas accounting for inter-individual differences in the N1b is actively avoided during GA-PD, due to the potential for the N1 amplitude to distort the N1b results, in the case of component selection its potential inclusion would only serve to improve the selection process. This is due to the fact that under the experimental conditions used for the visual task both the N1 and N1b evoked potentials will present at the level of the scalp with very similar and overlapping source topographies, this then leads to the neuronal activity for these evoked potentials to found within a single independent component.

Following component selection, the component activation waveform was back-projected for the channel of interest where manual peak detection was performed for each of the conditional averages for each participant. The back-projected ICA component peak latencies were then used to define mean amplitude time windows for the task-specific event-related SOIs and a measure of the SOI was calculated for the conditional averaged EEG data using these ICA-informed time windows. Essentially ICA-PD operates through

a very similar process to GA-PD, with each technique only differing in how the mean amplitude time windows were defined. As the ICA-PD technique individually tailors the SOI time windows for each participant it was hypothesised to outperform the GA-PD based signal analysis.

2.4.4. *Component of interest peak detection (COIPD)*

COIPD was performed using the same ICA decomposition and component selection methodologies as generated for the ICA-PD technique. The only difference between the two techniques being that, under COIPD, the amplitude of the SOI was calculated directly from the back-projected conditional average component activation waveforms for each participant. This technique was hypothesised to produce the best results as the back-projected ICA activity for the SOI should be well separated from other non-event-related neuronal activity and artefactual noise. Therefore, assessment of the SOI with COIPD should be less prone to measurement errors giving a more accurate account of task-specific event-related changes within the SOI.

Chapter 3. Passive Auditory Task (IDAEP)

3.1. Background

The brains electrical response to auditory stimuli has been a topic of interest for EEG researchers for the past 80 years, from the very infancy of the electroencephalographic technique right up until the present day. In the very first experiments, monopolar recordings at the vertex (position Cz of the international 10-20 positioning system), with a linked earlobes reference, uncovered the presence of an "on-effect" following the presentation of auditory stimuli. A biphasic and occasionally triphasic negative leading waveform that could be distinguished from the background neural data as an effect lasting approximately 0.3 seconds with a peak-to peak amplitude of around 100μV was initially described (Davis, 1939). The observation of this characteristic waveform in the raw data constitutes the very first publication of what is now referred to as the N1 and P2 stimulus evoked potentials. As technological advances were made in EEG and computing, the functional significance of the N1/P2 auditory ERP complex, or 'vertex-potential', was gradually uncovered. For example in 1965, it was the design and development of an on-line digital averaging computer (HAVOC, by Engebretson (1963)) that allowed Davis and Zerlin (1966) to show the amplitude sensitivity of the vertex-potential to the intensity of an auditory stimulus, as well as investigate the impact of stimulus frequency, rise and fall times and inter-stimulus-intervals (Davis and Zerlin, 1966). Studies such as this, where the strength of the auditory N1 and P2 components are associated with physical or sensory changes within the auditory environment, have led to these components being thought of as reflecting processes of selective attention and feature extraction within the auditory cortex (Paiva *et al.*, 2016). Although the N1 and P2 components are very often studied and referred to as a single entity or complex, the components themselves are thought to be independent processes that are generated within different cortical regions (Crowley and Colrain, 2004). Furthermore, the N1 component itself has actually been shown to represent the summation of multiple sub-components that are generated within different regions of the primary and secondary visual cortices (Naatanen and Picton, 1987; Woods, 1995; Zouridakis *et al.*, 1998; Paiva *et al.*, 2016). As the two components are not produced by the same neural generators many studies have tried to dissociate their activity into various distinct perceptual processes. However, a clear separation is yet to be achieved for all of the many different elements of selective attention and feature detection. For instance, although there is

good evidence to show that the N1 and P2 components demonstrate differential responses to stimulus duration (Alain *et al.*, 1997), with the N1 component showing sensitivity over a greater range of stimulus durations. There is, however, still conflicting evidence for how the N1 and P2 amplitudes relate to stimulus-intensity when levels exceed approximately 70dB SL (Adler and Adler, 1989; Hensch *et al.*, 2008).

A positive correlation between the N1 and P2 auditory evoked potential amplitudes and stimulus-intensity has been shown in many publications (Beagley and Knight, 1967; Picton *et al.*, 1970; Tepas *et al.*, 1972; Schweitzer and Tepas, 1974; Kaskey *et al.*, 1980; Bruneau *et al.*, 1985; Hensch *et al.*, 2008; Hagenmuller *et al.*, 2011). However, the specific nature of this relationship is still uncertain. Originally, Davis and Zerlin categorised the relationship as a power function with an exponent equal to that of the slope of a perceived linear banding of logarithmically scaled amplitude/stimulus-intensity datapoints. This interpretation of the data was noted to be in accordance with the power law that had been already observed for subjective audiometric tests of loudness/stimulus-intensity (Stevens and Guiroa, 1964). However, out of the seven participants tested by Davis and Zerlin, one was noted to exhibit unusually large vertex-potentials in response to the auditory stimuli and three others demonstrated amplitude stimulus-intensity functions (ASFs) that noticeably plateaued at higher intensities. Therefore, due to the wide amount of variation observed in such a small dataset, this early study likely lacked sufficient power to accurately assess the ASF relationship. Due to the difficulty of trying to assess the ASF relationship when such large amounts of variation existed in such small sample sizes, researchers began to consider the possibility that the ASF may be just as likely to actually be based upon a linear relationship as opposed to that of a power function (Davis *et al.*, 1968). Subsequent studies with greater participant numbers have since found the ASF to be best described by a linear relationship (Tepas *et al.*, 1972; Schweitzer and Tepas, 1974; Kaskey *et al.*, 1980). Whereby the strength of this relationship is calculated as the slope of a straight line fitted to the amplitude/intensity values by a mean least squares technique. However as mentioned early, there is also some evidence that this stimulus-intensity relationship may break down at higher stimulus-intensity levels (Bruneau *et al.*, 1985; Adler and Adler, 1989), as a result of a centrally descending overstimulation protection mechanism (Silverman *et al.*, 1969; Picton *et al.*, 1970; Zuckerman *et al.*, 1974; von Knorring *et al.*, 1978). Therefore, rather than assuming linearity some researchers have taken the approach of calculating the gradient of the ASF by more than one technique.

For example, Hegerl *et al.* (1992) estimated the gradient of the ASF by first using a LMS regression based approach, therefore assuming linearity. This was then followed by calculating the median slope of all possible connections between the different amplitudes, which as a robust statistic doesn't directly assume linearity. One benefit of using the median slope approach, that should be mentioned, is that it has been shown to be more reliable under test-retest conditions then other approaches (Hegerl *et al.*, 1992).

Irrespective of the methodology used to assess the amplitude/stimulus-intensity relationship, the discovery of its sensitivity to central serotonergic functioning (Hegerl and Juckel, 1993; Juckel *et al.*, 1999; Hegerl *et al.*, 2001; Kenemans and Kähkönen, 2011) has prompted its use in many clinical studies as a potential biomarker for disorders believed to be associated with serotonergic dysfunction (Linka *et al.*, 2007; Linka *et al.*, 2009; Lee *et al.*, 2012; Ostermann *et al.*, 2012; Wyss *et al.*, 2013; Juckel, 2015; Hagenmuller *et al.*, 2016). Furthermore, the ASF has been argued to be able to predict treatment efficacy in psychiatric disorders such as depression, bipolar and schizophrenia (Hegerl *et al.*, 1992; Juckel *et al.*, 2003; Linka *et al.*, 2004; Lee *et al.*, 2005). However, at present there is still no general consensus on the best way to isolate and assess the amplitude sensitive components of auditory evoked potentials (AEPs). With careful consideration and optimization of the AEP analysis pipeline it may therefore be possible to further improve the clinical predictive accuracy of the ASF measure beyond what has previously been observed. Therefore, it is the aim of this study to determine the optimal preprocessing parameters and the best signal analysis technique for assessing the ASF of an intensity-dependent auditory evoked potential (IDAEP) task. Any improvements in the estimation of the underlying signal will then help to better characterise the amplitude/stimulus-intensity relationship and consequently, any reduction in the variance of the estimated signal will aid statistical outcomes and therefore its clinical usefulness.

3.2. Methods

3.2.1. *Participants*

27 (17 female) right-handed participants with a mean age of 45.6 (SD 9.3) and no reported hearing impairment were recruited as healthy controls (HCs) for inclusion in a larger study of antiglucocorticoid augmentation of antidepressants in depression (ADD study, (McAllister-Williams *et al.*, 2016)). As a part of the ADD study, all participants were screened for a personal or familial (limited to first-degree relatives) history of mood disorders and identified themselves as being free from any medication with known neuropsychological effects. All participants provided written informed consent and were paid a small honorarium for their participation in the ADD study.

3.2.2. *Task*

The IDAEP task consisted of a single tonal auditory stimulus (1000 Hz sine wave, 50ms duration) being presented binaurally at 5 different intensity levels (60dB, 70dB, 80dB, 90dB and 100dB) through a pair of Sony MDR-CD270 digital reference headphones. 200 trials (40 per intensity level) were pseudo-randomly organised and presented at an inter-stimulus-interval (ISI) of 2050ms (+/-200ms). Participants were asked to sit quietly and focus on a fixation cross presented on a screen in front of them during the task and were informed that no response was required to be made and to simply let the sounds wash over them.

3.2.3. *EEG Recording*

During the task EEG was recorded using a Synamps2 (Neuroscan, Compumedics) amplifier with 34 Ag/AgCl electrodes (FPz, AFz, Fz, FCz, Cz, CPz, Pz, Oz, FP1, FP2, FC3,FC4, C3, C4, CP3, CP4, P3, P4, O1, O2, F7, F8, FT7,FT8, T7, T8, TP7, TP8, P7, P8, M1, M2) arranged by the international 10-20 system (Klem *et al.*, 1999), with a right mastoid reference (re-referenced to linked mastoids offline) and grounded at AFz. Two additional bipolar channels recorded HEOG and VEOG from 5 additional electrodes placed at the outer canthus of each eye, at the glabella and centrally below each lower eyelid. Data was recorded at a sampling rate of 500 Hz with a 0.1-100 Hz analogue band pass filter.

3.3. Analysis

Data preprocessing was performed in accordance with the preprocessing optimisation pipeline (POP) schematic detailed in Figure 2 (page 26) with the soft parameter settings for each configuration as detailed in Table 2. As is commonly observed with auditory ERP tasks with a linked mastoid reference the channel of interest was defined as Cz as it produced the strongest and clearest signal. Stimulus-locked epochs were made for all trials from -200ms to 1600ms. This allowed for the maximum amount of EEG data to be present for training the ocular artefact algorithms within the gamma POP configuration without any overlapping data between trials. All epochs were subsequently reduced in size (-200ms to 500ms post-stimulus) prior to baseline correction and the assessment and removal of epochs with voltages exceeding +/- 50 μV. As detailed in the analysis pipeline configurations chapter, participant exclusion for each analysis pipeline dataset was performed either based upon participants failing to have at least 20 epochs per condition or if any conditional average failed to meet the minimum baseline variability threshold criteria (7.7451μV^2 (4dp)). Once all participant data had been preprocessed for a given POP configuration, grand averages of the signal and noise indexes (BV, SV and SNR) were then calculated for that dataset to allow for later comparison with the other POP configurations.

The N1-P2 mean amplitude difference was assessed through calculation of N1 and P2 peak amplitudes via the different SAP methodologies detailed in section 1.2 earlier. From this data the gradient of the ASF was calculated in three ways. The median slope of all 14 of the possible connections between the five amplitude/stimulus-intensity datapoints was calculated for each participant to allow for easy comparison of the data with other research groups that favour this methodology. Following this the same data was also fitted to linear and polynomial mixed effects models to allow for a more direct assessment of the ASF relationship. Within these models, linear and polynomial fixed effects were used to characterise the slope and a participant grouping variable was used to define a random intercept. The mixed effects models were conducted using the *fitlme* function within Matlab with the following Wilkinson notation for each model:

Linear Model: Amplitude ~ Intensity + (1|Participant)

Polynomial Model: Amplitude ~ Intensity^2 + (1|Participant)

Constant Model: Amplitude ~ 1 + (1|Participant)

The linearity of the data was assessed by a likelihood ratio test between the linear mixed effects model and the constant model. The validity of adding a polynomial variable to the model was assessed via a likelihood ratio test between the linear and polynomial models. The impact of the different POP and SAP configurations upon accuracy of the linear and polynomial model based ASF gradients were assessed through changes in their r^2. Statistical analysis of the median slope defined ASF gradient simply focused upon a one sample t-test against a population mean of zero and calculation of the Cohen's d effect size. For the purpose of creating a smooth narrative within this thesis only the linear and polynomial results and statistics will be presented in the results section of this chapter. However, for completeness the statistical results based upon the median slope defined ASF gradient measure have been provided within the appendix section of this thesis (Chapter 6.2 , page 127).

Table 2. IDAEP POP soft parameter configurations. Where no filter cut-off value is present in the table the amplifier filter settings at the time of recording apply (i.e. LFP 0.1Hz and HPF 100Hz)

Configuration Number	Configuration Type	Clean Line Noise	HPF Cutoff (Hz)	LPF Cutoff (Hz)	OAR Type
01	Raw	×	×	×	×
02	Alpha	✓	×	×	×
03	Beta	✓	0.5	×	×
04	Beta	✓	1.0	×	×
05	Beta	✓	×	30	×
06	Beta	✓	×	60	×
07	Beta	✓	0.5	30	×
08	Beta	✓	0.5	60	×
09	Beta	✓	1.0	30	×
10	Beta	✓	1.0	60	×
11	Gamma	✓	1.0	30	LMS
12	Gamma	✓	1.0	30	CRLS
13	Gamma	✓	0.5	30	LMS
14	Gamma	✓	0.5	30	CRLS
15	Gamma	✓	×	30	LMS
16	Gamma	✓	×	30	CRLS

3.4. Results

3.4.1. *BV exclusion criteria threshold and PVAF window limits*

Out of all of the data preprocessed by the wild-type POP variant, configuration 11 with a 1.0Hz HPF, 30Hz LPF and ocular artefact removal (OAR) based upon LMS regression was found to produce the best grand average SNR (Table 3). As expected the auditory evoked potential responses of each stimulus-intensity type could clearly be seen within the grand averaged conditional data of channel Cz (Figure 4). The wild-type data demonstrated the characteristic N1 and P2 peak amplitude sensitivity to stimulus-intensity and the ASF of N1-P2 mean amplitude difference was also observed to positively correlate with stimulus-intensity (Figure 5).

Using the GA-PD signal analysis approach on the optimal POP configuration data (Config. 11), the BV exclusion criteria threshold was calculated and set at 7.7451μV^2 (4dp). This value was determined based upon the SOI being defined as the amplitude change associated with a 10dB increase in stimulus-intensity (i.e. the smallest stimulus-intensity step change tested within the IDAEP task). This was therefore calculated by multiplying the grand average intensity gradient of the linear model (0.2783 (4dp)) by the stimulus-intensity change of the SOI (10) and then squaring the value to give the BV_{max} exclusion criteria threshold. It is noteworthy that due to an unforeseen oversight in the analysis process the BV_{max} exclusion threshold was not set using a median-based approach as originally intended. Therefore, the calculation used is not as robust to potential biasing from outlying participant data that may be present within the wild-type data variant.

The same optimised dataset and conditional grand average waveforms were also used to set the PVAF window limits for the ICA-based signal analysis at 125ms-256.5ms, based upon the average N1 (145ms) and P2 (236.5ms) peak latencies across all of stimulus intensities with the addition of +/-20ms to the window to allow for inter-individual differences.

Table 3. Wild-type POP variant signal and noise index measures and GA-PD linear model results for the likelihood ratio test between the linear and constant models (L vs C) and the coefficient of determination for the linear model (r^2). The best SNR was observed for configuration 11 (1.0Hz HPF, 30Hz LPF, OAR by LMS). Highly significant L vs C likelihood ratio test results were observed across all datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model. POP configuration 9 (1.0Hz HPF, 30Hz LPF, no OAR) produced the largest ASF linear effect size of all the datasets tested. POP configuration 11 produced the largest ASF linear effect size of the POP configuration datasets that included OAR (Gamma types).

POP Configuration						Linear Model	
Number	Type	N	BV	SV	SNR	L vs C p-value	r^2
1	Raw	27	3.2008	26.2101	0.0798	0.00E+00	0.7122
2	Alpha	27	3.1525	26.1672	0.0803	0.00E+00	0.7041
3	Beta	27	3.1510	24.1829	0.0869	0.00E+00	0.7042
4	Beta	27	3.1186	21.4408	0.1008	0.00E+00	0.7139
5	Beta	27	2.5636	25.5322	0.0909	0.00E+00	0.7055
6	Beta	27	2.8524	25.8710	0.0857	0.00E+00	0.7004
7	Beta	27	2.5628	23.5483	0.0998	0.00E+00	0.7052
8	Beta	27	2.8507	23.8868	0.0932	0.00E+00	0.7002
9	Beta	27	2.5297	20.8024	0.1209	0.00E+00	0.7145
10	Beta	27	2.8186	21.1440	0.1104	0.00E+00	0.7096
11	Gamma	27	2.2259	19.3717	0.1545	0.00E+00	0.7137
12	Gamma	27	2.2371	19.3357	0.1541	0.00E+00	0.7128
13	Gamma	27	2.2491	21.5626	0.1368	0.00E+00	0.7108
14	Gamma	27	2.2513	21.8649	0.1343	0.00E+00	0.7129

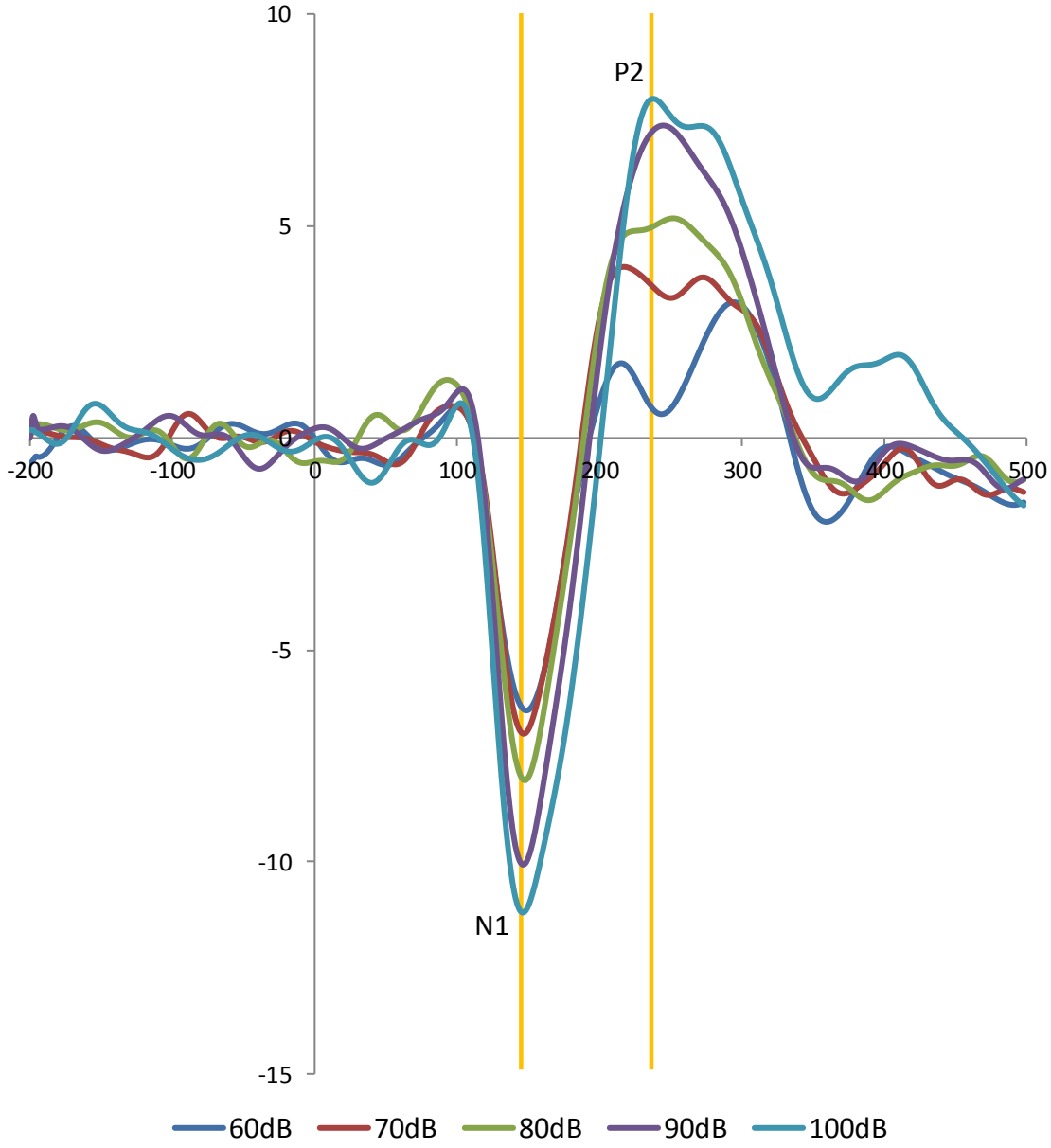

Figure 4. Wild-type grand average ERP waveform for Config.11 (IDAEP task) where the ordinate relates to amplitude (μV) and the abscissa relates to time (ms). Yellow lines represent the average N1 (145ms) and P2 (236.5ms) peak latencies of the stimulus-intensity evoked responses. The wild-type grand average conditional peak latency values that were used to inform the N1 and P2 individual mean amplitude measurements were N1: 148ms, 144ms, 146ms, 144ms, 146ms; P2: 216ms, 216ms, 250ms, 242ms, 238ms (for the 60dB, 70dB, 80dB, 90dB and 100dB intensity auditory stimuli respectively).

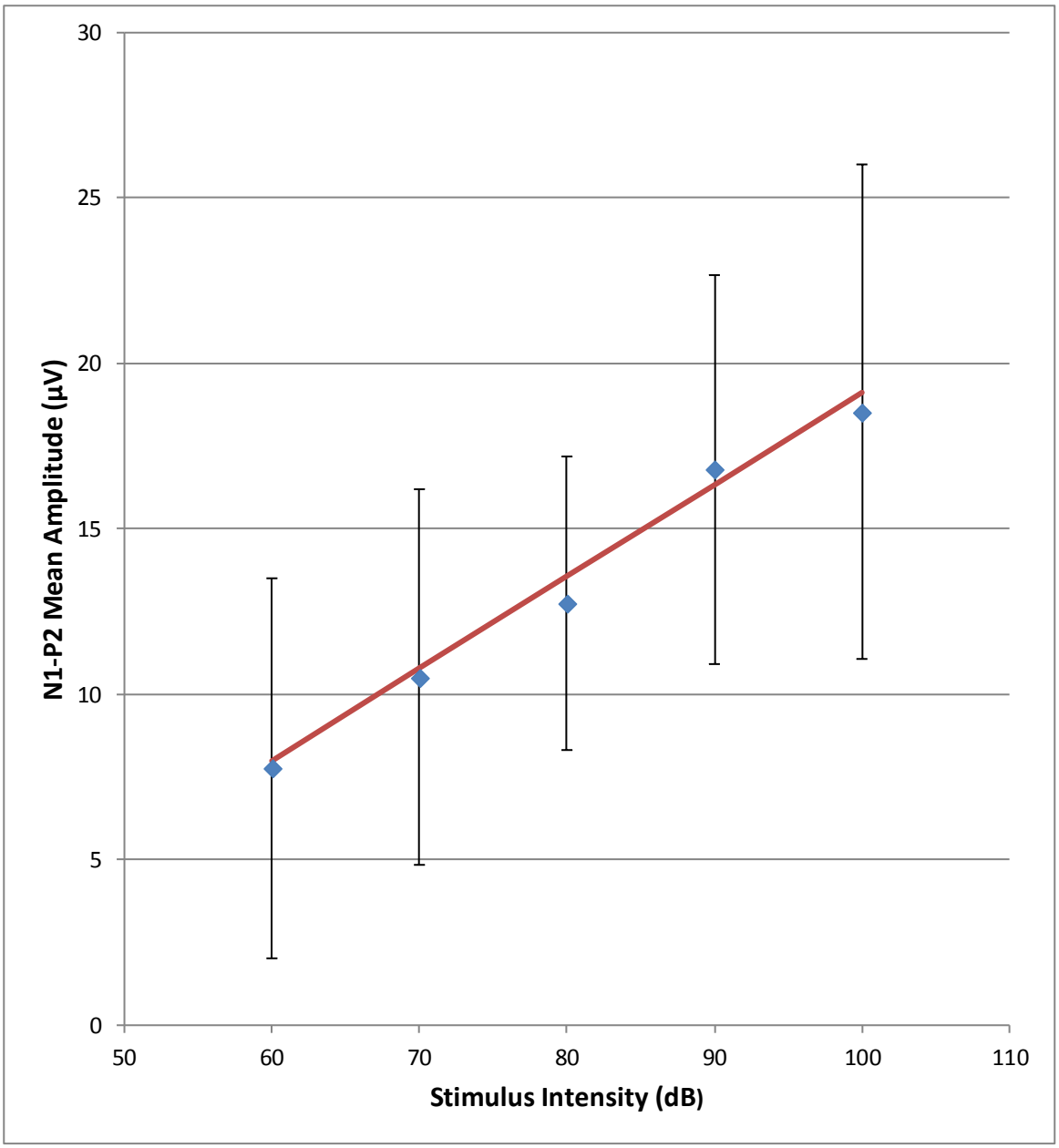

Figure 5. Wild-type ASF for Config.11 using GA-PD (IDAEP task) for calculation of N1-P2 mean amplitude difference for each stimulus-intensity. Error bars relate to standard deviation and the intact line represents the linear ASF model with a gradient of 0.2783 (4dp).

3.4.2. *POP output*

When data preprocessing was performed under normal conditions (i.e. not using the wild POP variant) and included data pruning via the exclusion of epochs with excessive voltage potentials and participant exclusion based upon either of the two data quality exclusion criteria, sub-optimal POP configurations with poor artefact correction processes could be seen to suffer from large numbers of participants being excluded from the dataset (Table 4 and Table 5). Therefore, to ensure that the conclusions made about the different POP and SAP configurations were representative of a population estimate rather than simply a participant sub-group, interpretation of the signal and noise indexes (BVs, SVs and SNRs) for the different configurations were only conducted for datasets were at least half of the participants remained following participant exclusion (i.e. when N>27/2). However, as none of the datasets from the beta type POP configurations displayed sufficient participant numbers following participant exclusion under the minimum epoch number threshold exclusion approach (ExCrit.Epoch$_{min}$), the decision of which filter settings to carry forward to the gamma type POP configurations was made by selecting a range of filter settings based upon those datasets that had the highest number of participants remaining (i.e. POP configurations 5, 7 & 9). Under the maximum BV exclusion criterion (ExCrit.BV$_{max}$) approach both configurations 7 and 9 each possessed sufficient participant numbers for selection of the gamma filtering parameters to be made based upon that which produced the best SNR. However, rather than restricting analysis to just that of the filter parameters of configuration 9 (1.0Hz HPF, 30Hz LPF), it was decided to use the same range of filter settings as for the ExCrit.Epoch$_{min}$ analysis pipeline. This then allowed for easier comparison of the two different exclusion criteria methodologies as well as allowing for assessment of a broader range of filter parameters which also happen to relate to the best performing three beta-type POP configuration SNRs for the ExCrit.BV$_{max}$ datasets.

Regardless of the method of participant exclusion, gamma-type pipeline configuration 11 (1.0Hz HPF, 30Hz LPF, OAR by LMS regression) was seen to produce the best SNRs (ExCrit.Epoch$_{min}$ SNR=0.1843; ExCrit.BV$_{max}$ SNR=0.1708). Conversely, the best BVs and SVs were all seen for different configurations (ExCrit.Epoch$_{min}$ BV=2.2054 (Config.16), SV=22.6550 (Config.15); ExCrit.BV$_{max}$ BV=1.7427 (Config.21), SV=22.2545 (Config.15)).

Table 4. Pruned POP signal and noise index measures following participant exclusion via the minimum epoch number exclusion threshold criterion approach (IDAEP task). POP configuration 11 (1.0Hz HPF, 30Hz LPF, OAR by LMS) produced the best SNR out of the datasets where the number of participants remaining (N) following participant exclusion was greater than 13 (i.e. were at least half of the original 27 participants remained following participant exclusion).

POP Configuration		Epoch.min Exclusion Threshold			
Number	Type	N	BV	SV	SNR
1	Raw	1	3.4513	30.2040	0.1693
2	Alpha	1	3.3262	30.8923	0.1720
3	Beta	1	2.9701	27.4454	0.1794
4	Beta	1	2.8701	24.7146	0.2028
5	Beta	7	1.6014	23.9043	0.1813
6	Beta	5	2.0841	24.1394	0.1670
7	Beta	9	2.7249	22.0976	0.1761
8	Beta	5	2.1509	22.3735	0.1722
9	Beta	10	3.3952	22.3901	0.1929
10	Beta	5	2.1763	19.9873	0.1908
11	Gamma	24	2.4470	18.9187	0.1708
12	Gamma	25	2.4298	18.4905	0.1694
13	Gamma	24	2.5154	20.8863	0.1545
14	Gamma	25	2.4212	20.6288	0.1570
15	Gamma	20	2.7484	22.6550	0.1515
16	Gamma	21	2.2054	22.3408	0.1550

Table 5. Pruned POP signal and noise index measures following participant exclusion via the maximum BV exclusion threshold criterion (IDAEP task). POP configuration 9 (1.0Hz HPF, 30Hz LPF, no OAR) produced the best SNR out of all the beta type datasets tested. POP configuration 11 produced the best SNR of the POP configuration datasets that included OAR (Gamma types).

POP Configuration		BV.max Exclusion Threshold			
Number	Type	N	BV	SV	SNR
1	Raw	4	3.1976	22.5517	0.1613
2	Alpha	4	3.1771	18.4570	0.1312
3	Beta	3	2.8215	16.0166	0.1333
4	Beta	4	2.8950	15.9486	0.1413
5	Beta	12	2.0732	21.7871	0.1569
6	Beta	7	2.2338	21.1961	0.1452
7	Beta	15	2.2482	21.5167	0.1671
8	Beta	7	2.4340	15.7725	0.1324
9	Beta	15	2.6968	20.3785	0.1826
10	Beta	8	2.7449	17.9736	0.1612
11	Gamma	21	1.84390	18.1195	0.1843
12	Gamma	21	1.7427	17.7676	0.1814
13	Gamma	21	1.8450	20.8448	0.1681
14	Gamma	21	1.7552	20.1057	0.1671
15	Gamma	19	1.9684	22.2545	0.1593
16	Gamma	23	2.1494	21.7188	0.1507

3.4.3. *GA-PD*

From this point on the same interpretation restriction as above, i.e. requiring a necessary minimum number of participants per dataset (N>27/2), was used for all further analyses reported in this chapter. For all datasets and POP configurations incorporating either types of participant exclusion, the likelihood ratio tests comparing the linear and constant ASF models for GA-PD signal analysis were all found to be significantly in favour of a linear model (Table 6 and Table 7). Furthermore, the addition of a polynomial factor did not significantly improve the fit of the model for any dataset. POP configuration 13 (HPF 0.5, LPF 30, OAR by LMS regression) was found to produce the linear model with the largest ASF effect size in both exclusion criteria data groups (ExCrit.Epoch$_{min}$ r^2=0.7675; ExCrit.BV$_{max}$ r^2=0.7465).

Table 6. Pruned GA-PD ExCrit.Epoch$_{min}$ linear and polynomial model statistics (IDAEP task). With likelihood ratio tests between constant (C), linear (L) and polynomial (P) models. Highly significant L vs C likelihood ratio test results were observed across all datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model for these datasets. No significant P vs L likelihood ratio test results were observed for any of the datasets where at least half of the original 27 participants remained following participant exclusion. POP configuration 13 (0.5Hz HPF, 30Hz LPF, OAR by LMS) produced the largest ASF linear effect size of the datasets where at least half of the original 27 participants remained following participant exclusion.

POP Configuration			Linear Model		Polynomial Model		
Number	Type	N	L vs C p-value	r^2	P vs C p-value	P vs L p-value	r^2
1	Raw	1	5.30E-04	0.9094	1.12E-04	1.29E-02	0.9737
2	Alpha	1	5.38E-04	0.9089	9.43E-05	1.04E-02	0.9755
3	Beta	1	9.97E-04	0.8854	2.16E-04	1.40E-02	0.9659
4	Beta	1	4.08E-03	0.8079	8.40E-04	1.50E-02	0.9411
5	Beta	7	1.95E-12	0.8836	1.67E-11	7.55E-01	0.8840
6	Beta	5	8.09E-08	0.8497	1.58E-07	1.12E-01	0.8688
7	Beta	9	1.52E-12	0.8295	6.98E-12	2.44E-01	0.8364
8	Beta	5	4.48E-07	0.8157	1.52E-06	2.52E-01	0.8284
9	Beta	10	2.39E-12	0.8219	1.57E-11	4.30E-01	0.8249
10	Beta	5	2.52E-07	0.8355	3.86E-07	8.61E-02	0.8597
11	Gamma	24	0.00E+00	0.7498	0.00E+00	9.77E-01	0.7498
12	Gamma	25	0.00E+00	0.7299	0.00E+00	8.67E-01	0.7300
13	Gamma	24	0.00E+00	0.7675	0.00E+00	7.86E-01	0.7677
14	Gamma	25	0.00E+00	0.7350	0.00E+00	9.14E-01	0.7350
15	Gamma	20	1.11E-16	0.7447	6.66E-16	6.85E-01	0.7453
16	Gamma	21	0.00E+00	0.7189	0.00E+00	6.32E-01	0.7198

Table 7. Pruned GA-PD ExCrit.BV_{max} linear and polynomial model statistics (IDAEP task). With likelihood ratio tests between constant (C), linear (L) and polynomial (P) models. Highly significant L vs C likelihood ratio test results were observed across all datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model for these datasets. No significant P vs L likelihood ratio test results were observed for any of the datasets. POP configuration 13 (0.5Hz HPF, 30Hz LPF, OAR by LMS) produced the largest ASF linear effect size of the datasets where at least half of the original 27 participants remained following participant exclusion.

POP Configuration			Linear Model		Polynomial Model		
			L vs C		P vs C	P vs L	
Number	Type	N	p-value	r^2	p-value	p-value	r^2
1	Raw	4	9.29E-07	0.8068	5.36E-06	6.54E-01	0.8095
2	Alpha	4	2.54E-05	0.7293	1.16E-04	5.31E-01	0.7367
3	Beta	3	6.80E-04	0.7644	3.09E-03	9.05E-01	0.7647
4	Beta	4	1.04E-03	0.6233	4.32E-03	7.12E-01	0.6270
5	Beta	12	6.40E-12	0.7268	5.61E-11	9.61E-01	0.7268
6	Beta	7	6.88E-08	0.7660	3.05E-07	3.41E-01	0.7743
7	Beta	15	3.39E-11	0.7030	2.56E-10	6.32E-01	0.7043
8	Beta	7	2.89E-06	0.6958	1.50E-05	5.67E-01	0.6998
9	Beta	15	1.38E-10	0.6562	1.13E-09	9.64E-01	0.6562
10	Beta	8	5.88E-07	0.6806	3.74E-06	8.39E-01	0.6811
11	Gamma	21	0.00E+00	0.7287	0.00E+00	7.23E-01	0.7292
12	Gamma	21	0.00E+00	0.7436	0.00E+00	6.29E-01	0.7444
13	Gamma	21	0.00E+00	0.7465	0.00E+00	6.24E-01	0.7473
14	Gamma	21	0.00E+00	0.7387	0.00E+00	4.31E-01	0.7409
15	Gamma	19	1.11E-15	0.7200	1.12E-14	8.19E-01	0.7202
16	Gamma	23	0.00E+00	0.6934	0.00E+00	4.34E-01	0.6957

3.4.4. *IAPD*

As for the GA-PD data, comparison of the goodness of fit of the IAPD datasets to either a constant or linear model significantly favoured a linear model (Table 8 and Table 9). However, two ExCrit.Epoch$_{min}$ POP configuration datasets (Configs. 11 and 12) displayed marginally significant results for the polynomial/linear likelihood ratio test. Suggesting that the ASFs of these specific datasets may be best described by a polynomial modal (Poly.Vs.Lin: p=5.67E-02, p=5.83E-02 respectively). The POP configuration that produced the largest effect size was different for each of the two participant exclusion methodology groups (ExCrit.Epoch$_{min}$ r^2=0.7827(Config.11); ExCrit.BV$_{max}$ r^2=0.7901(Config.13)). Additionally, both effect sizes were an improvement upon the best GA-PD results for either exclusion criteria group (Δ Effect Size: ExCrit.Epoch$_{min}$ +1.99% (2dp), ExCrit.BV$_{max}$ +5.84% (2dp)).

Table 8. IAPD ExCrit.Epoch$_{min}$ linear and polynomial model statistics (IDAEP task). With likelihood ratio tests between constant (C), linear (L) and polynomial (P) models. Highly significant L vs C likelihood ratio test results were observed across all datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model for these datasets. Marginally significant P vs L likelihood ratio tests results were observed for two of the datasets where at least half of the original 27 participants remained following participant exclusion. POP configuration 11 (1.0Hz HPF, 30Hz LPF, OAR by LMS) produced the largest ASF linear effect size of the datasets where at least half of the original 27 participants remained following participant exclusion.

POP Configuration			Linear Model		Polynomial Model		
Number	Type	N	L vs C p-value	r^2	P vs C p-value	P vs L p-value	r^2
1	Raw	1	5.88E-04	0.9058	1.70E-04	1.85E-02	0.9690
2	Alpha	1	8.11E-04	0.8939	2.46E-04	2.01E-02	0.9640
3	Beta	1	5.97E-04	0.9053	1.04E-04	1.04E-02	0.9745
4	Beta	1	2.35E-03	0.8428	1.31E-04	3.32E-03	0.9720
5	Beta	7	1.62E-09	0.8011	1.26E-08	9.37E-01	0.8012
6	Beta	5	4.08E-06	0.7659	1.53E-05	3.29E-01	0.7780
7	Beta	9	2.38E-11	0.7865	1.96E-10	7.79E-01	0.7871
8	Beta	5	1.67E-05	0.7510	8.42E-05	6.35E-01	0.7541
9	Beta	10	9.81E-12	0.7899	5.07E-11	3.07E-01	0.7959
10	Beta	5	1.20E-05	0.7657	5.34E-05	4.72E-01	0.7723
11	Gamma	24	0.00E+00	0.7827	0.00E+00	5.67E-02	0.7916
12	Gamma	25	0.00E+00	0.7719	0.00E+00	5.83E-02	0.7808
13	Gamma	24	0.00E+00	0.7803	0.00E+00	2.45E-01	0.7837
14	Gamma	25	0.00E+00	0.7721	0.00E+00	1.48E-01	0.7773
15	Gamma	20	0.00E+00	0.7798	0.00E+00	1.75E-01	0.7853
16	Gamma	21	0.00E+00	0.7650	0.00E+00	1.13E-01	0.7727

Table 9. IAPD ExCrit.BV_{max} linear and polynomial model statistics (IDAEP task). With likelihood ratio tests between constant (C), linear (L) and polynomial (P) models. Highly significant L vs C likelihood ratio test results were observed across all datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model for these datasets. No significant P vs L likelihood ratio tests results were observed for any of the datasets. POP configuration 13 (0.5Hz HPF, 30Hz LPF, OAR by LMS) produced the largest ASF linear effect size of the datasets where at least half of the original 27 participants remained following participant exclusion.

POP Configuration			Linear Model		Polynomial Model		
			L vs C		P vs C	P vs L	
Number	Type	N	p-value	r^2	p-value	p-value	r^2
1	Raw	4	5.12E-06	0.7942	3.04E-05	9.18E-01	0.7944
2	Alpha	4	3.36E-05	0.7701	1.63E-04	6.26E-01	0.7738
3	Beta	3	1.64E-04	0.8496	6.92E-04	5.54E-01	0.8543
4	Beta	4	3.49E-05	0.8453	3.55E-05	6.69E-02	0.8765
5	Beta	12	2.19E-11	0.7309	1.80E-10	7.79E-01	0.7314
6	Beta	7	3.03E-06	0.7051	1.38E-05	4.46E-01	0.7120
7	Beta	15	2.01E-12	0.7448	1.43E-11	4.91E-01	0.7471
8	Beta	7	2.50E-05	0.6929	1.28E-04	6.84E-01	0.6949
9	Beta	15	2.24E-11	0.7069	1.86E-10	8.01E-01	0.7072
10	Beta	8	1.09E-06	0.6615	5.37E-06	4.76E-01	0.6676
11	Gamma	21	0.00E+00	0.7775	0.00E+00	1.90E-01	0.7825
12	Gamma	21	0.00E+00	0.7707	0.00E+00	1.90E-01	0.7758
13	Gamma	21	0.00E+00	0.7901	0.00E+00	2.87E-01	0.7933
14	Gamma	21	0.00E+00	0.7742	0.00E+00	2.95E-01	0.7775
15	Gamma	19	0.00E+00	0.7883	0.00E+00	6.68E-01	0.7888
16	Gamma	23	0.00E+00	0.7506	0.00E+00	2.33E-01	0.7549

3.4.5. *ICA-PD*

Due to the fact that ICA decomposition is both a computationally intensive and lengthy process, the effectiveness of the different ICA-based signal analysis techniques were only tested on a subset of pipeline configuration datasets (Configs. 11, 13 and 15). The specific configurations were chosen as they had consistently produced the largest effect sizes in the GA-PD and IAPD analysis stages. Using ICA-PD all ASF data in this subset was seen to significantly fit to a linear model when compared to a constant model, with no significant improvement with the addition of a polynomial factor (Table 10 and Table 11). The largest effect sizes were seen under POP configuration 15 for both of the exclusion criteria datasets (r^2: ExCrit.Epoch$_{min}$ 0.7371, ExCrit.BV$_{max}$ 0.7539). Under ExCritBV$_{max}$ configuration 15 the ICA-PD technique produced a larger effect size than when the same data was analysed by GA-PD (Δ Effect Size: 4.71 (2dp)). However, for all other configurations and datasets ICA-PD was seen to be inferior to GA-PD signal analysis.

Table 10. ICA-PD ExCrit.Epoch$_{min}$ linear and polynomial model statistics (IDAEP task). With likelihood ratio tests between constant (C), linear (L) and polynomial (P) models. Of the limited number of POP configurations analysed by an ICA-based approach highly significant L vs C likelihood ratio test results were observed for all three datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model for these datasets. No significant P vs L likelihood ratio tests results were observed for any of three datasets. POP configuration 15 (No digital HPF, 30Hz LPF, OAR by LMS) produced the largest ASF linear effect size of three datasets.

POP Configuration			Linear Model		Polynomial Model		
Number	Type	N	L vs C p-value	r^2	P vs C p-value	P vs L p-value	r^2
11	Gamma	24	0.00E+00	0.7293	0.00E+00	1.11E-01	0.7373
13	Gamma	24	0.00E+00	0.7032	0.00E+00	2.95E-01	0.7069
15	Gamma	20	0.00E+00	0.7371	0.00E+00	1.43E-01	0.7449

Table 11. ICA-PD ExCrit.BV_{max} linear and polynomial model statistics (IDAEP task). With likelihood ratio tests between constant (C), linear (L) and polynomial (P) models. Of the limited number of POP configurations analysed by an ICA-based approach highly significant L vs C likelihood ratio test results were observed for all three datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model for these datasets. No significant P vs L likelihood ratio tests results were observed for any of three datasets. POP configuration 15 (No digital HPF, 30Hz LPF, OAR by LMS) produced the largest ASF linear effect size of three datasets.

POP Configuration			Linear Model		Polynomial Model		
Number	Type	N	L vs C p-value	r^2	P vs C p-value	P vs L p-value	r^2
11	Gamma	21	0.00E+00	0.7152	0.00E+00	3.11E-01	0.7191
13	Gamma	21	0.00E+00	0.7048	0.00E+00	4.27E-01	0.7073
15	Gamma	19	0.00E+00	0.7539	0.00E+00	8.34E-01	0.7541

3.4.6. *COIPD*

Similar to ICA-PD, COIPD was seen to produce data that was significantly represented by a linear model (Table 12 and Table 13), however for configurations 13 and 15, of both exclusion criteria groups, the likelihood ratio tests showed that a polynomial model produced a significantly better fit for these configurations than that given by a linear model (Poly.Vs.Lin: ExCrit.$Epoch_{min}$ p=1.08E-02 (Config.11), p=4.33E-02 (Config.13); ExCrit.BV_{max} p=1.90E-02 (Config.11), p=2.54E-02 (Config.13)).

However, the largest effect size for the ExCrit.BV_{max} group was not seen for either of these more complex models, instead signal analysis favoured the linear model of configuration 15 (r^2=0.8235). The largest effect size for the ExCrit.$Epoch_{min}$ group was also seen for configuration 15 (r^2=0.7279). Comparing the best significantly linear ASF effect size results from this technique (i.e. Config. 15) with the best significantly linear ASF effect size results obtained from GA-PD (i.e. Config 13), preprocessing utilising ExCrit.$Epoch_{min}$ produced a weaker effect size using COIPD signal analysis compared to GA-PD (Δ Effect Size: -5.16% (2dp)). However, when preprocessing utilized ExCrit.BV_{max} COIPD produced a stronger effect size than the best ExCrit.BV_{max} GA-PD analysis (Δ Effect Size: +10.31% (2dp)).

Table 12. COIPD ExCrit.Epoch$_{min}$ linear and polynomial model statistics (IDAEP task). With likelihood ratio tests between constant (C), linear (L) and polynomial (P) models. Of the limited number of POP configurations analysed by COIPD, highly significant L vs C likelihood ratio test results were observed for all three datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model for these datasets. Significant P vs L likelihood ratio tests results were observed for two of the three datasets indicating that the ASF for these results were better characterised by a polynomial rather than linear relationship. POP configuration 15 (No digital HPF, 30Hz LPF, OAR by LMS) produced the largest ASF linear effect size of three datasets, which was also greater than either of the effect sizes for the significant polynomial model results.

POP Configuration			Linear Model		Polynomial Model		
Number	Type	N	L vs C p-value	r^2	P vs C p-value	P vs L p-value	r^2
11	Gamma	24	1.13E-12	0.6292	3.99E-13	1.08E-02	0.6571
13	Gamma	24	1.68E-14	0.6725	2.12E-14	4.33E-02	0.6880
15	Gamma	20	1.95E-11	0.7279	8.64E-11	2.51E-01	0.7329

Table 13. COIPD ExCrit.BV$_{max}$ linear and polynomial model statistics. With likelihood ratio tests between constant (C), linear (L) and polynomial (P) models. Of the limited number of POP configurations analysed by COIPD, highly significant L vs C likelihood ratio test results were observed for all three datasets indicating that the ASF was better represented by a sloped linear relationship than a constant model for these datasets. Significant P vs L likelihood ratio tests results were observed for two of the three datasets indicating that the ASF for these results were better characterised by a polynomial rather than linear relationship. POP configuration 15 (No digital HPF, 30Hz LPF, OAR by LMS) produced the largest ASF linear effect size of the three datasets analysed with COIPD ExCrit.BV$_{max}$ as well as the largest effect size in comparison to all other signal analysis methods, exclusion methodologies and POP configurations.

POP Configuration			Linear Model		Polynomial Model		
Number	Type	N	L vs C p-value	r^2	P vs C p-value	P vs L p-value	r^2
11	Gamma	21	9.44E-12	0.5960	5.27E-12	1.90E-02	0.6255
13	Gamma	21	1.22E-15	0.6884	9.99E-16	2.54E-02	0.7088
15	Gamma	19	3.89E-15	0.8236	3.79E-14	7.78E-01	0.8237

3.5. Conclusions

The preprocessing optimisation pipeline was designed to encourage the systematic exploration of artefact rejection techniques. Through the analysis of wild-type data it has been possible to see how each technique in the pipeline gradually improves the overall quality of the data and how each step can be refined through adjustment of various parameter settings. The improvements in SV and SNR both significantly correlated with the effect size of the wild GA-PD analysis (Figure 6). However, none of the observed index measures of signal quality or noise estimation could accurately predict which configuration would produce the largest effect size. Not being able to predict this at the end of the preprocessing optimisation stage, i.e. before signal analysis, somewhat limits the general usability of the pipeline optimisation approach. Therefore, for a researcher to be confident that they have achieved the optimal level of preprocessing for subsequent signal analysis they would need to follow a potentially lengthy process of signal analysis for a selection of the best performing parameter configurations. The limited usefulness of the BV, SV and SNR indexes may be a product of their overgeneralised calculation and non-specificity. For instance, simply taking the average BVs, SVs and SNRs across all conditions and participants ignores a wealth a data relating to the distribution, spread and range of inter and intra individual differences for each of the indexes. Consideration of this may help to better understand the quality of the dataset as a whole and improve its predictive capabilities. Further improvements may also be possible for the SV and SNR indexes through adjustment of the time window used for their calculation. The broad post-stimulus potential time window settings (i.e. 0ms to 500ms) used here may be better replaced with ones that specifically focus upon the timeframe of the ERP components of interest (i.e. using the same time window as the PVAF calculation).

Although improving the predictive power of the optimisation outcome measures would be welcomed, it could be argued that the importance of preprocessing optimisation and even to some extent preprocessing in general to directly improving the effect size is somewhat questionable. Comparison of the various preprocessing configurations was initially performed using a wild-type variant of the POP (i.e. with no epoch or participant exclusions) to ensure that the noise reducing benefits inherent in the averaging process applied equally across all of the configurations, therefore, allowing for an unbiased assessment of the configurations. Under these conditions the actual impact of the

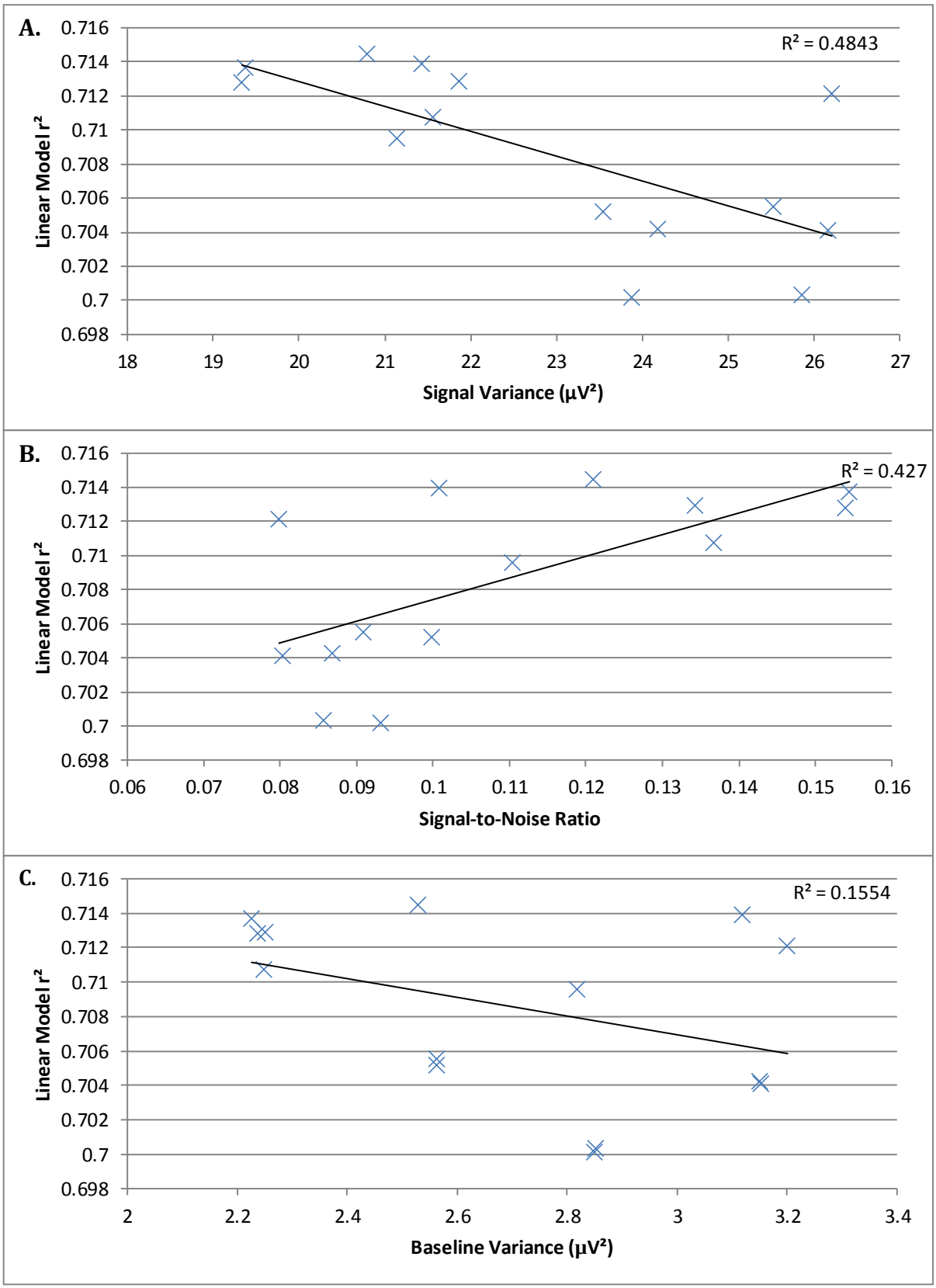

Figure 6. Wild GA-PD POP linear model effect size (r^2) for each configuration and their relation to A) Signal Variance (SV), B) Signal-to-noise ratio (SNR) and C) Baseline Variance (BV) (IDAEP task). R^2 value for the Pearson correlation is displayed with highly significant correlations observed for SV and SNR (p=0.0047 and p=0.0097 respectively)

various preprocessing configurations was fairly negligible. The moderate but significant correlations of SV and SNR with the effect size of the wild-type GA-PD analysis only related to a change of effect size across a very small range ($GA\text{-}PD_{Wild}$: r^2 Range=0.0143). Therefore, in reality the impact of the different preprocessing optimisation configurations was somewhat arbitrary when the data was analysed using the wild-type POP variant. The limited effect of the various preprocessing techniques highlights the effectiveness of the averaging process as a noise reduction technique. GA-PD analysis benefits from averaging at two different stages, firstly when all of the epochs are used to make the conditional averages for individual participants and secondly when all of the participants are used to make the grand average. The large amount of data that was guaranteed by the wild-type POP variant allowed the averaging process to create a sufficiently accurate estimate of the population signal. Peak detection performed on the grand average was then used to set appropriate time windows for calculating the N1 and P2 mean amplitudes for each participant and condition. Even with no preprocessing, with 40 epochs per condition and 20 participants, it has been possible to demonstrate a significantly linear ASF with an effect size that was greater for wild-type GA-PD data than for some of the more complicated pruned variant POP and SAP configurations. However, it is worth noting that the success of a wild-type analysis pipeline approach will be highly dependent upon having a dataset that is largely of 'good quality'. If a substantial portion of the data was made up of participants with excessive levels of non-ergodic noise (i.e. noise that may be time or phase locked to the stimulus event) then the averaging process would not be able to resolve this and the results would be biased or masked by this noise. Therefore, even though the data produced by the IDAEP task was able to produce useful and interpretable results with no preprocessing or participant exclusion it could be argued that solely following a wild-variant pipeline analysis approach should be avoided in favour of assessing and removing potentially damaging epochs and participants from the analysis pipeline, as is common practice in most EEG research.

The standard process of pruning epochs/trials that contain excessive levels of noise is an overly conservative means of artefact removal as the process itself rejects data based upon consideration of the voltage potentials in all of the EEG channels across the entire scalp rather than restricting itself to just the region or channel of interest specific to the task. In theory artefact rejection is performed in this way as even though an artefact might not be maximal within the task-specific region of interest the signal of a large

artefact will actually propagate across the scalp and affect all channels to some degree (unless the artefact is due to poor electrode contact with the scalp). However, this process can sometimes lead to a dramatic loss of data for a given participant and therefore requires the researcher to establish participant exclusion criteria to account for these situations where very little 'clean' data remains. As detailed in the methodology section the process of classifying participant datasets for exclusion has been approached using two different methodologies. The common method of setting a minimum epoch number threshold per condition (ExCrit.$Epoch_{min}$) and a new approach involving direct assessment of the noise levels present in the remaining data and ensuring it doesn't exceed a maximum threshold level (ExCrit.BV_{max}). The indirect and insensitive nature of the $Epoch_{min}$ threshold approach for participant data quality assurance was evident in both the results of the earlier/non-optimised and later/optimised, pruned POP configurations. During the non-optimized stages, unsuccessful artefact removal typically resulted in a mass rejection of data and therefore under the $Epoch_{min}$ threshold approach any participant that didn't have at least 20 epochs in each test condition (i.e. the 5 different stimulus intensity levels) was believed to have data of a questionable quality and was therefore excluded. However, when the noise in the remaining data is directly assessed via consideration of the baseline variance at the channel of interest relevant for the task (i.e. Cz in the case of the IDAEP task) a greater proportion of the participants were found to be acceptable for further analysis (Table 6 and Table 7). At the later stages of POP optimisation, the inadequacy of the $Epoch_{min}$ threshold approach was evident in its occasional over inclusivity of participants with poor data quality. Furthermore, within these POP configuration datasets a few participants with 'sufficient' epoch numbers were actually removed from analysis under the BV_{max} participant exclusion approach as regardless of meeting the epoch number threshold the data was still found to contain a considerable amount of noise.

The composition and various strengths of all the different noise artefacts that can affect an EEG will be different for each participant. Therefore, the effectiveness of each of the different POP configurations for artefact removal will also be different for each participant. During the earlier POP configurations (i.e. before OAR is incorporated in the POP or the filter settings are optimised) very few participants made it past the exclusion criteria stage. As the vast majority of participants were excluded during these earlier pre-optimised POP configurations, this unfortunately removed any real validity in

assessing how each of the different techniques independently impacted upon the post-preprocessing signal and noise estimates or the ASF effect sizes of the different signal analysis approaches. Therefore, in generating the guideline that interpretation of the pipeline configurations must be restricted to only those where the majority of participants had passed the exclusion process, it unfortunately limited the scope of analysis to just a small subset of techniques and parameters. Furthermore, due to the ICA-based signal analysis approaches being both computationally and time intensive only datasets from three of the gamma type POP configurations where analysed by the COIPD and ICA-PD approaches and therefore comparative analysis was restricted to techniques and parameters involved in these same POP configurations (Configs. 11, 13 and 15).

After consideration of the signal analysis results from the 4 different approaches, for all the data from the 3 gamma type POP configuration subsets, it was clear that no one preprocessing configuration consistently produced the best results (Table 14). However, based upon the parameter make-up of the 3 configurations within the gamma subset, two processing techniques regularly appear to be beneficial to analysis outcomes, low-pass filtering and the ocular artefact removal technique. Low-pass filtering with a 30Hz cut-off vastly improved the number of participants in the preprocessed dataset. The advantage of this particular filter setting could be related to the successful removal of muscle artefacts that may be present in some of the data or it could be an indicator of inadequacies in the *Cleanline* function to be able to remove 50Hz line noise. In the case of ocular artefact removal, the LMS regression approach was seen to result in ASFs with larger effect sizes than those produced by CRLS regression, for all but one of the many POP and SAP configuration comparisons (GA-PD ExCrit.BV_{max} 11 and 12).

Table 14. Gamma type POP configuration subset effect sizes for each signal analysis technique (IDAEP task). Colour coding relates to effectiveness of the technique for a given dataset from best to worst, i.e. largest to smallest (green > yellow > orange > red)

	Config.	N	GA-PD	IAPD	ICA-PD	COIPD
$Epoch_{min}$	**11**	24	0.750	0.783	0.729	0.629
	13	24	0.767	0.780	0.703	0.672
	15	20	0.745	0.780	0.737	0.728
BV_{max}	**11**	21	0.729	0.778	0.715	0.596
	13	21	0.747	0.790	0.705	0.688
	15	19	0.720	0.788	0.754	0.824

Table 15. N1 and P2 peak latency statistics for ExCrit.BV_{max} Config.15 (IDAEP task)

Stimulus Intensity (dB)	N1 Peak Latency (ms)			P2 Peak Latency (ms)		
	GA-PD	IAPD Mean (St.Dev)	COIPD Mean (St.Dev)	GA-PD	IAPD Mean (St.Dev)	COIPD Mean (St.Dev)
60	150	150.86 11.10	149.62 17.67	212	269.62 40.33	240.38 37.80
70	146	147.71 9.57	146.57 11.80	214	244.29 37.18	234.00 35.82
80	146	143.24 12.97	146.76 22.79	256	244.00 31.90	240.38 26.90
90	144	144.67 8.43	148.38 16.21	236	252.19 28.16	239.05 29.00
100	146	150.38 11.47	147.81 23.05	238	256.67 30.90	245.90 31.81

Whilst the different POP configurations can be thought of as a means to improving the ASF results through the betterment of the artefact removal processes, the 4 different SAPs on the other hand are intended to affect the outcome of the task through improving the accuracy with which the signal itself is measured. Unlike the wild-type GA-PD analysis the various pruned POP and SAP configurations gave rise to a fairly broad range of results (Table 14). When simply considering the three main optimised preprocessing configurations (Configs. 11, 13 and 15), signal analysis with IAPD clearly stood out as the most consistent technique for producing the best ASF effect sizes. Furthermore the IAPD approach seemed to be relatively insensitive to the different POP configurations with very similar results being obtained irrespective of the HPF cut-off value or ocular artefact removal methodology used (Table 8 and Table 9). As expected, the GA-PD approach did not perform as well as IAPD, with a slight reduction observed in the ASF effect sizes for ExCrit.$Epoch_{min}$ data and a greater reduction for ExCrit.BV_{max} data. However rather than this being due to the failings of an 'over' generalised approach to peak amplitude calculation as hypothesised at the outset of the study, the failings of GA-PD may in fact be due to an underestimation of inter-individual differences. A brief investigation of the individual peak latencies from one of the IADEP datasets suggests that averaging across a time-window of 20ms for both the N1 and P2 peaks may be too narrow to account for the actual inter-individual differences in the data (Table 15). Based upon the fact that for a Gaussian process 95% of the data should fit within +/- 2 standard deviations from the mean, then given the IAPD latency statistics, a suitable window size for the N1 and P2 mean amplitude calculations might be ~40ms and ~70ms respectively.

Interestingly both of the ICA-based signal analysis techniques also failed to perform as expected, with the ICA-PD and COIPD approaches both performing worse than all of the other techniques for all but one dataset (ExCrit.BV_{max} Config.15). As the PVAF time window used for component selection was based upon the peak latencies of the wild GA-PD variant with the addition of +/- 20ms either side, the PVAF measure will have also been too small to capture the inter-individual P2 peak latency differences. However, as the N1 peak is well captured by the window and both peaks are expected to be found within a single component, the PVAF time window inaccuracies cannot be to blame for the poor performance of the ICA-based techniques.

For ICA-PD the calculation of the mean peak amplitudes for the N1 and P2 of each participant should theoretically be more accurate then under the generalised GA-PD approach as all conditions have the peak time windows customised for each participant. Although the impact and appropriateness of the size of the peak time window used in this instance in not fully understood, this factor alone could not account for the technique performing worse than GA-PD. Where the results show this to be true it suggests that either the ICA-driven characterisation of the N1 and P2 peaks was somehow inaccurate or that the true underlying ASF relationship is actually not well represented by a linear model. The fact that a significant polynomial relationship was observed for two datasets under COIPD lends some weight to the later explanation of the poor linear effect size results with ICA-PD. However, in further investigating the low intensity plateau which appeared to be driving this polynomial relationship (Figure 7) a potential confound relating to classification of the P2 peak was uncovered.

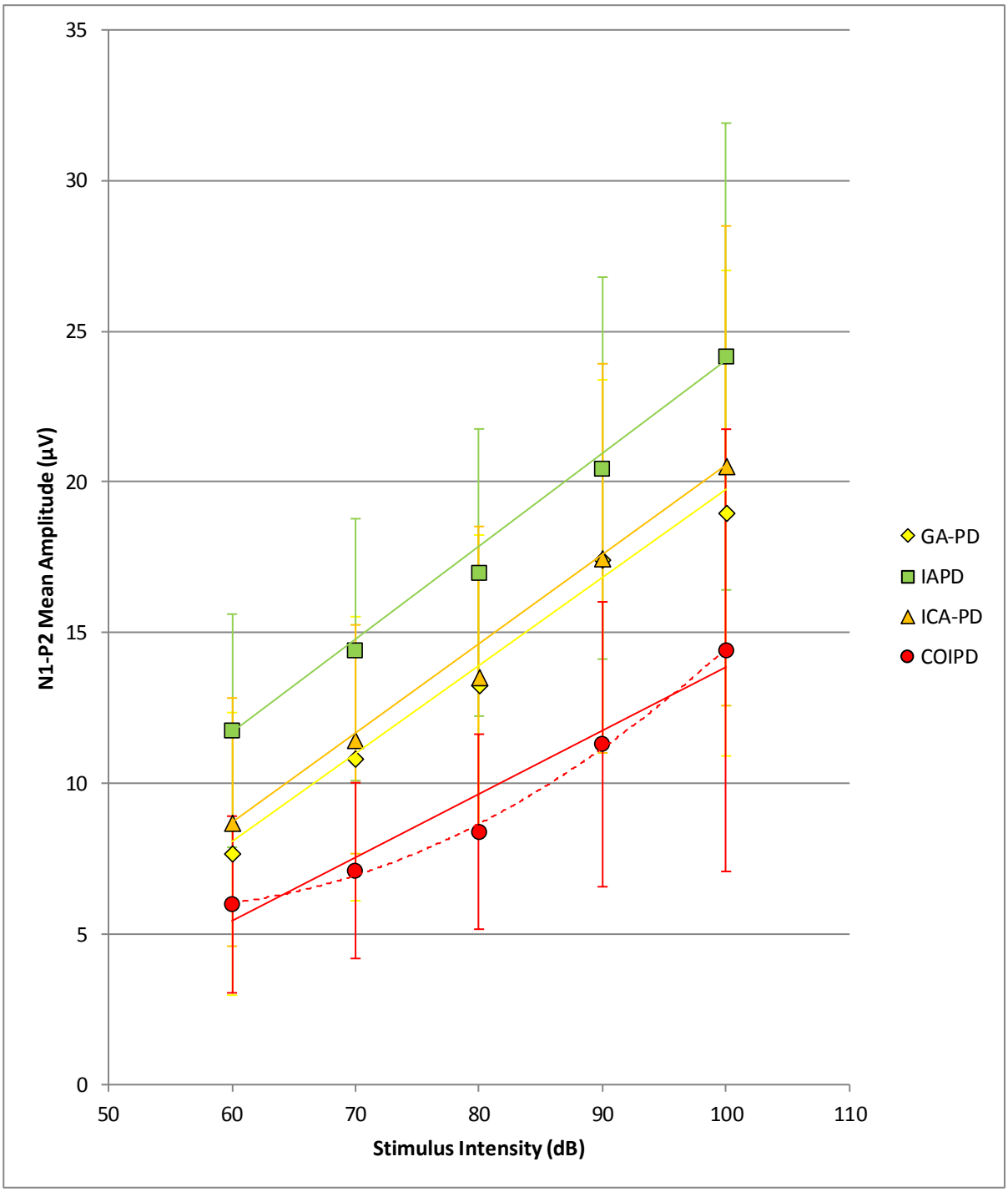

Figure 7. ASFs for ExCrit.BV$_{max}$ Config.13 (IDAEP task) where error bars relate to standard deviation and the intact and dashed lines indicate significant linear or polynomial ASF models respectively. Note how the gradient of linear model for COIPD ASF is shallower than all of the other signal analysis approaches and how this data also significantly fits to a polynomial model. The unique ASF characteristics of this COIPD dataset appears to be driven by the N1-P2 mean amplitude differences for the two lowest intensity auditory stimuli. The auditory evoked responses for these stimuli appear to be larger than would be expected if the data were to fit to the same gradient of linear relationship that appears to be present within the three largest intensity auditory stimuli, which is also of a similar gradient to the linear ASFs of the other signal analysis approaches.

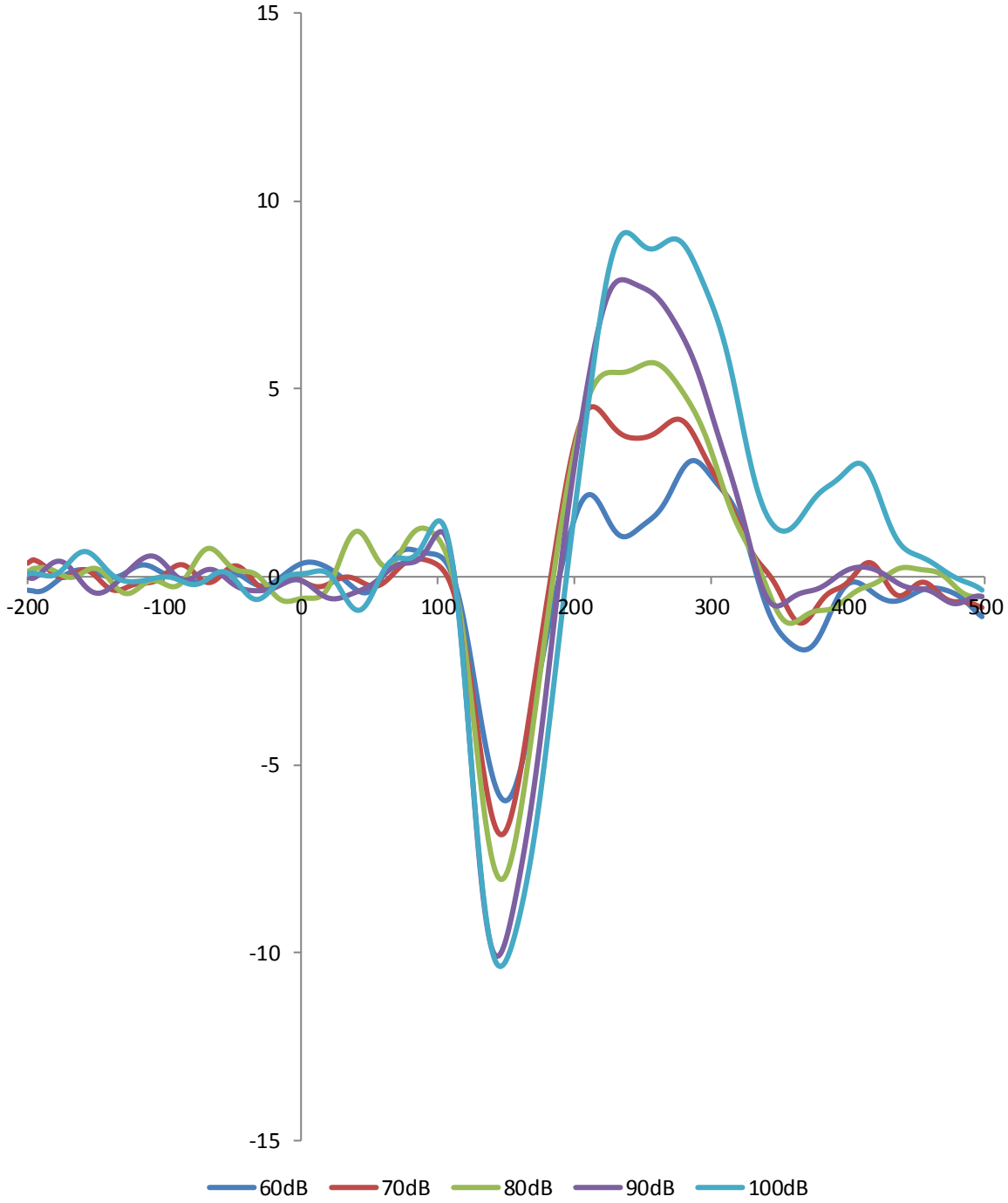

Figure 8. Grand average ERP waveform for ExCrit.BV$_{max}$ Config.13 (IDAEP task) where the ordinate relates to amplitude (μV) and the abscissa relates to time (ms). Note how the P2 ERP component appears to split into two smaller peaks for the lower intensity auditory stimuli.

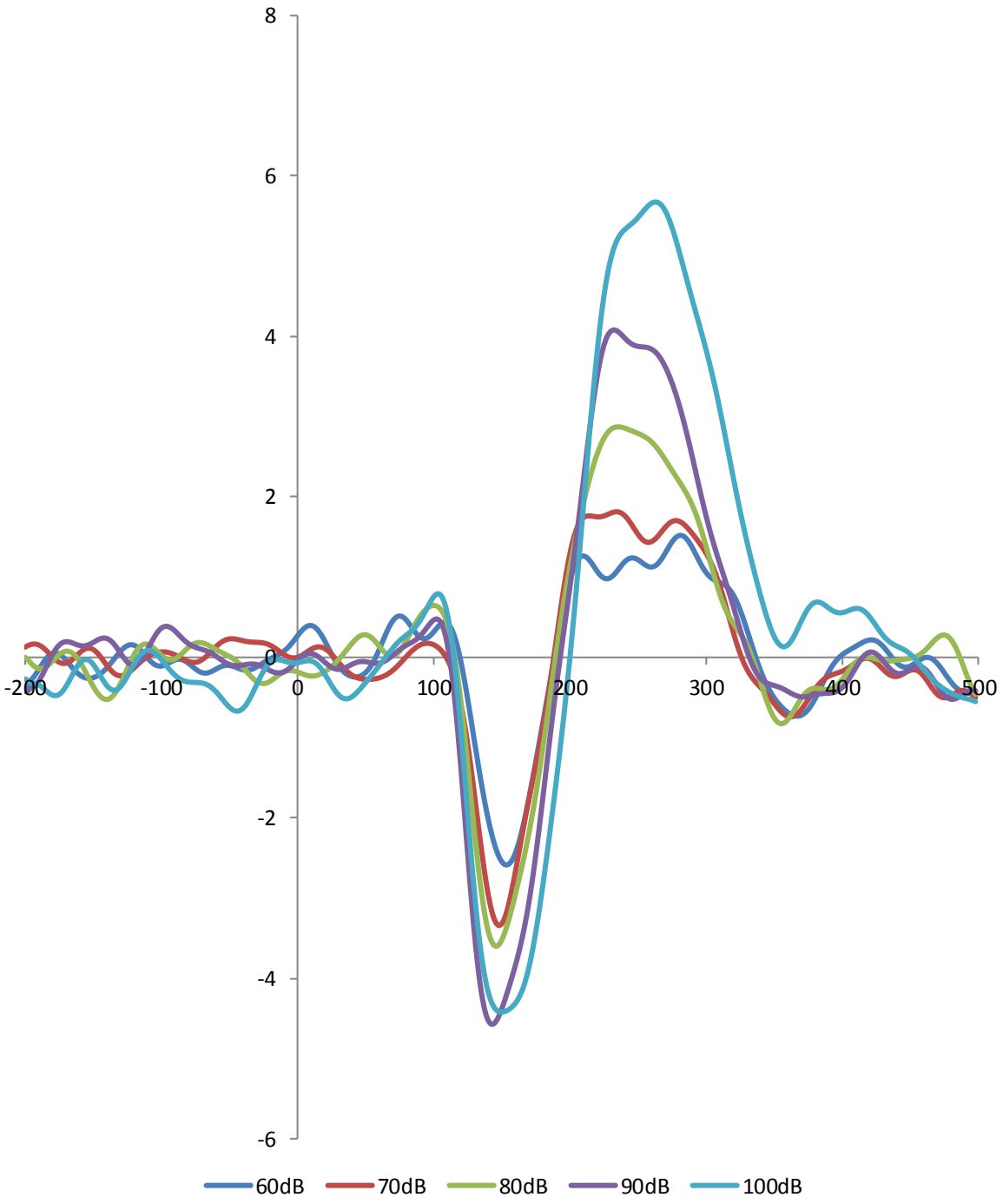

Figure 9. Grand average COI waveform for ExCrit.BV$_{max}$ Config.13 (IDAEP task) Note how the P2 component becomes somewhat ambiguous within the general P2 timeframe for the lower intensity auditory stimuli.

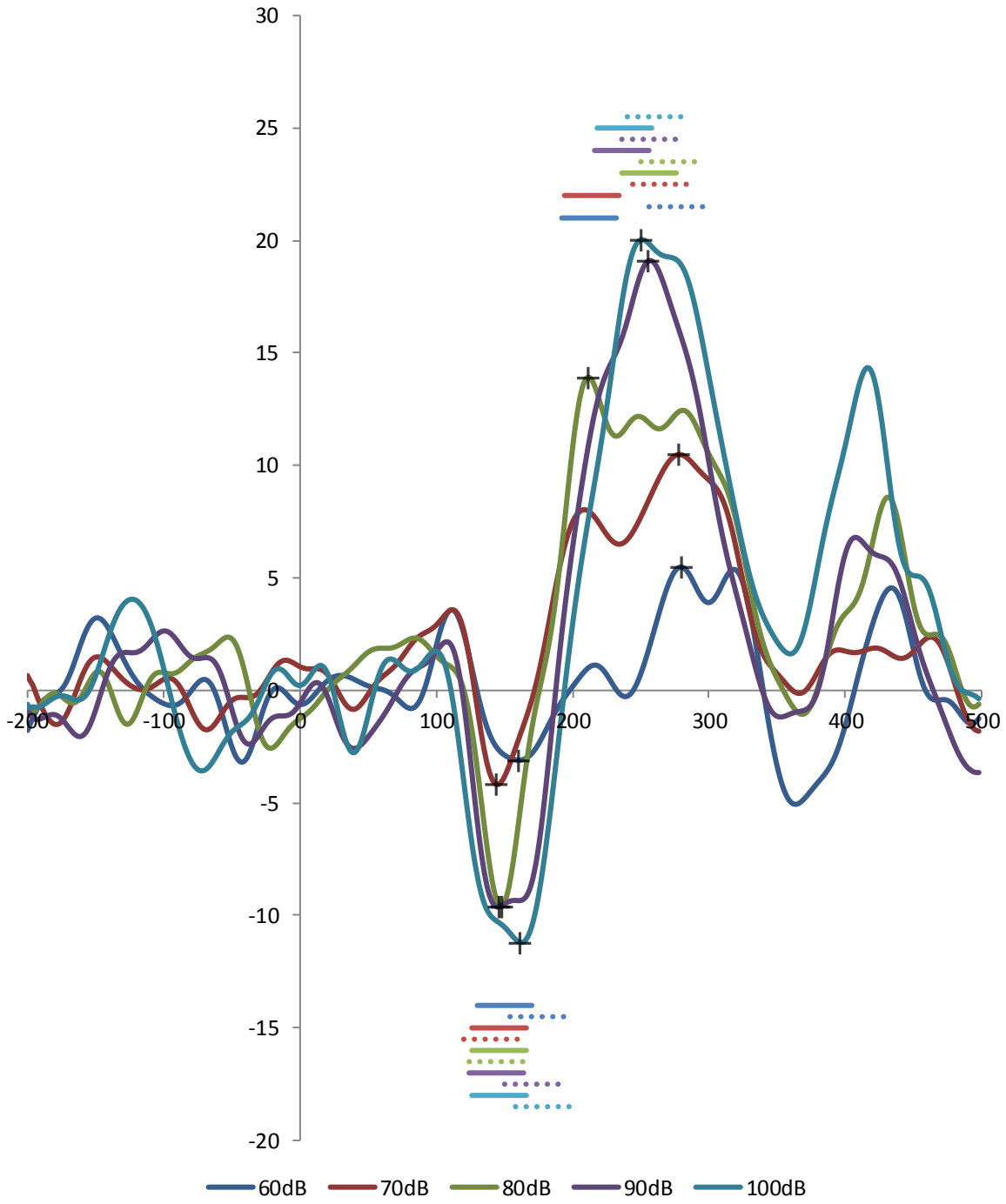

Figure 10. Participant 1 ERP average waveform ExCrit.BV$_{max}$ Config.13 (IDAEP task) where solid lines above and below the waveform denote the GA-PD N1 and P2 mean amplitude time windows. Dashed lines correspond with ICA-PD mean amplitude time windows. Cross markers relate to IAPD peak detection. Note how the GA-PD technique focuses upon an entirely different timeframe within the ERP waveform for its measurement of the P2 peak amplitude for the 60dB and 70dB stimulus-intensity conditions, compared to the other two techniques.

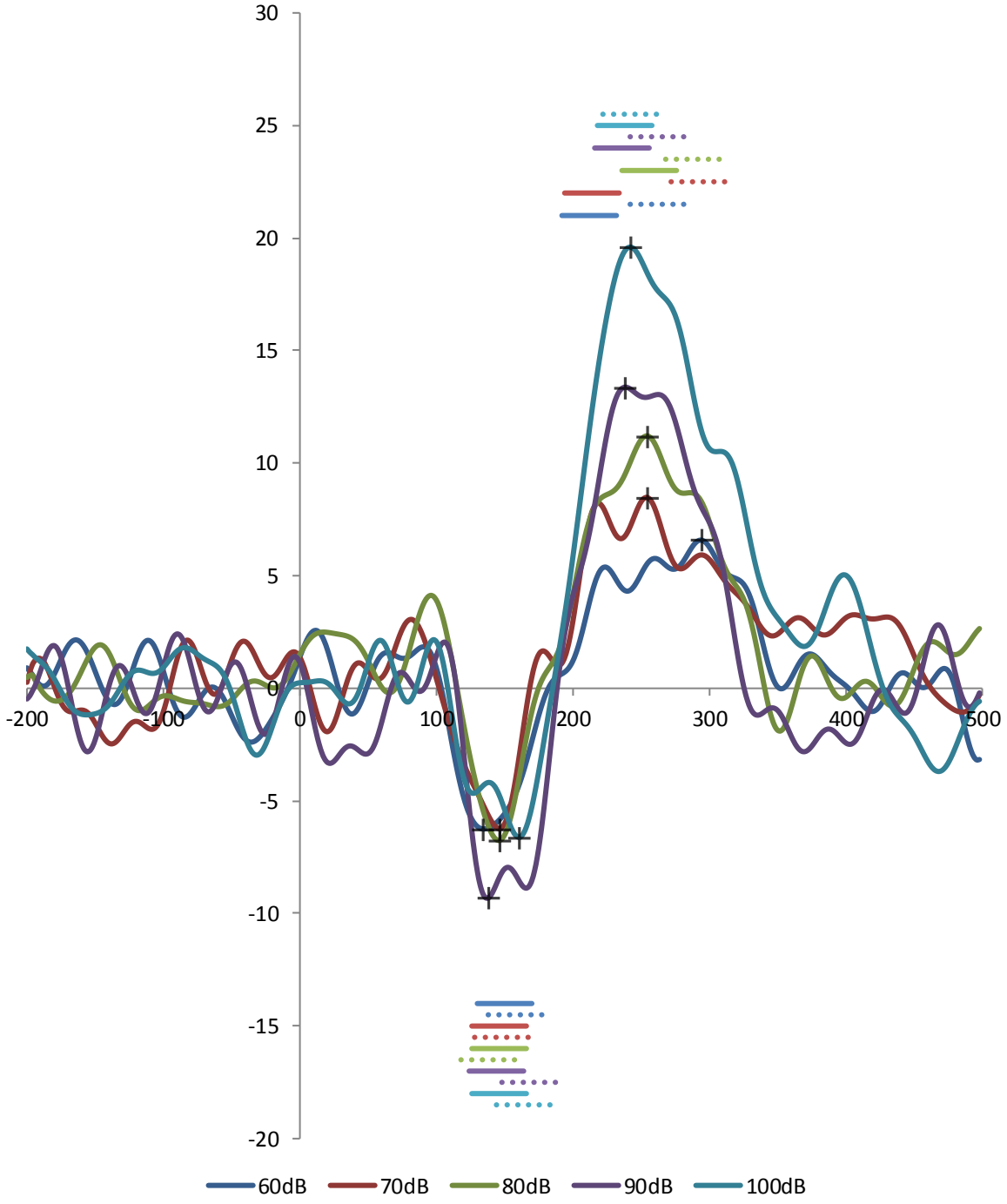

Figure 11. Participant 27 ERP average waveform ExCrit.BV$_{max}$ Config.13 (IDAEP task). Note how the GA-PD technique focuses upon an entirely different timeframe within the ERP waveform for its measurement of the P2 peak amplitude for the 60dB and 70dB stimulus-intensity conditions, compared to the other two techniques.

During N1 and P2 peak classification for many of the datasets it was noted that classification of the P2 peak was at times a little ambiguous for the low intensity stimuli. This ambiguity appeared to be caused by the P2 component splitting into two separate peaks for the lower intensity stimuli (Figure 8 and Figure 9). A re-examination of the IAPD and COIPD participant waveforms and the corresponding peak latency data highlighted how earlier peaks were often disregarded as noise with later peaks typically being favoured for their larger amplitudes (Figure 10 and Figure 11). Therefore, the improvement in effect size seen for the IAPD technique as compared to the GADP technique may in part be down to this shift in focus to later latencies for the low intensity stimuli. The presence of a split P2 component has not been observed or documented in the majority of IADP publications, however as many of the publications often fail to show either a grand average or an example participant waveform this may be a more widely experienced feature that is often overlooked. For example, Lee *et al.* (2005) performed peak detection for each participant through an automated process which simply defined the N1 and P2 peaks as the most negative and positive peaks within a predefined time window (N1: 65-175ms and P2: 120-280ms). By utilising a fully-automated peak detection process this would have meant that should any P2 peak splitting/ambiguity be present within the data the larger of the two peaks would always have been selected. However, in the published data for this study P2 peak splitting can be seen in one of the 80dB stimulus-intensity grand average waveforms however the larger of the two peaks hasn't been identified (Figure 12). This demonstrates how P2 peak splitting/ambiguity can be overlooked, as clearly peak detection on the published grand average waveform figure was performed by the researcher selecting the earlier of the two peaks rather than acknowledging and applying the same conditions used by the automated approach. Therefore, it is recommended that whenever automatic peak detection is performed with a broad detection time-window that statistics related to the peak latencies should be performed to check for peak splitting or simply to ascertain the inter-individual peak latency variation as an index of classification accuracy/peak stability. The presence and potential impact of P2 peak splitting upon the results of the ICA-based signal analysis approaches requires further analysis to ascertain whether the failure of ICA-PD and COIPD under configurations 11 and 13 is linked to P2 peak classification ambiguity (Figure 13 and Figure 14). Furthermore, it is unclear if the unique nature, i.e. producing the largest effect size of all tested POP and SAP variations, of ExCrit.BV_{max} Config.15 is related to the P2 peak splitting that can be observed within

the COI grand average (Figure 15), or if such a large improvement in effect size could be generated as the result of simply switching three participants and excluding one participant as occurs when comparing the two different exclusion criteria datasets for POP configuration 15 (Table 14).

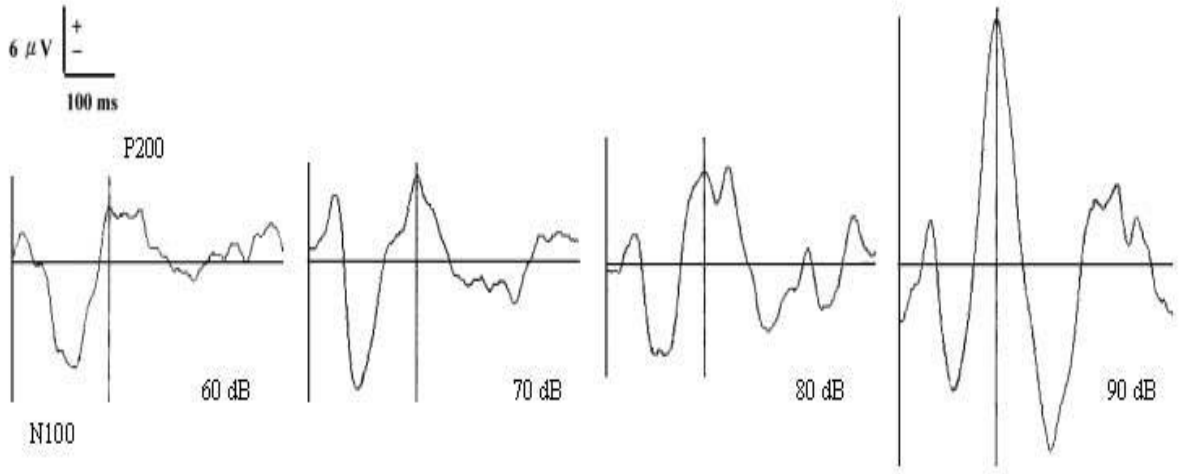

Figure 12. Grand average ERP waveforms demonstrating the auditory evoked potentials for four different stimulus intensities in a group of strong responders. Copied from Lee *et al.* (2005). Note how P2 peak splitting is present in the 80dB condition and P2 ambiguity can observed in the 60dB condition.

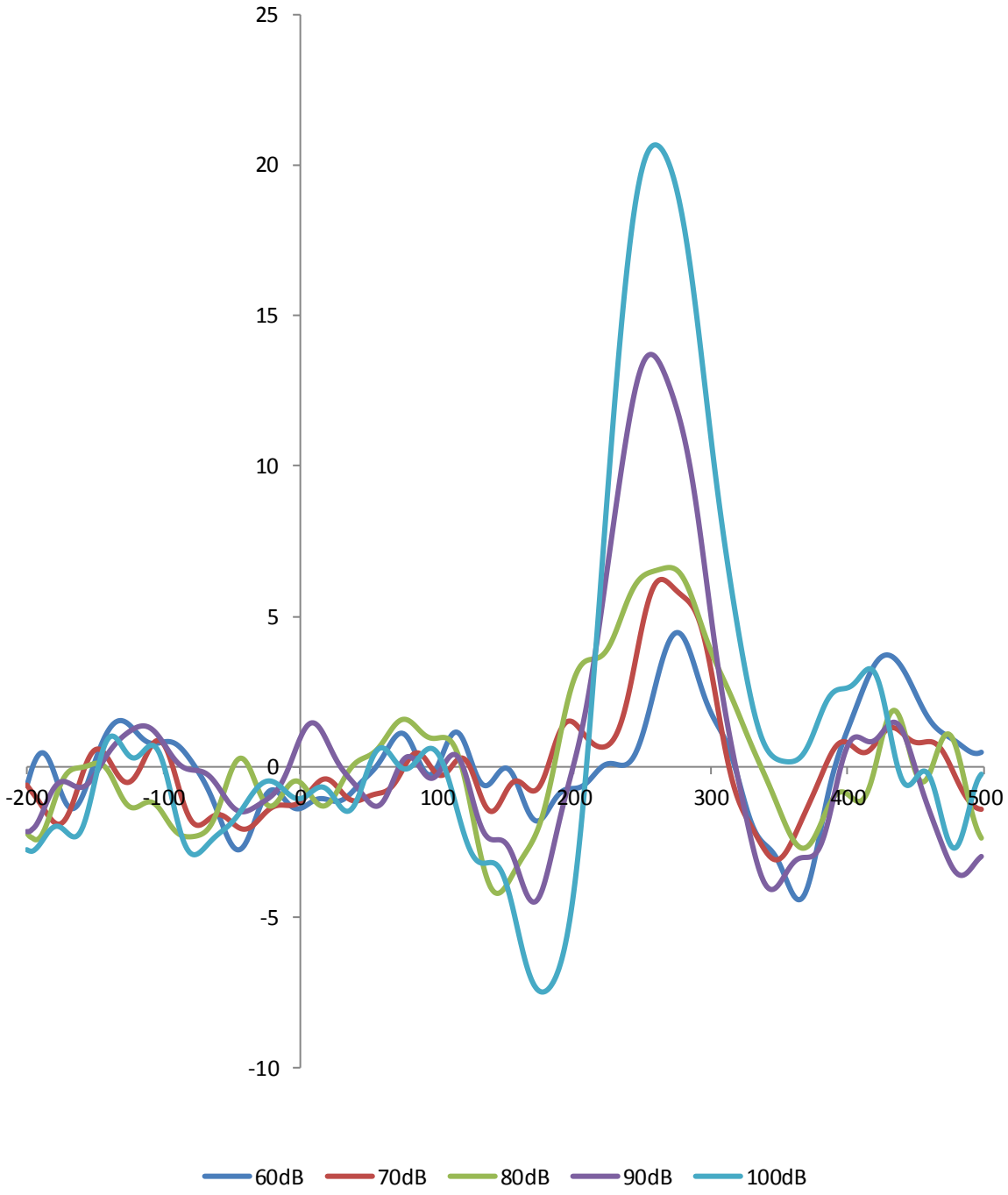

Figure 13. Participant 1 COI back-projected average waveform ExCrit.BV$_{max}$ Config.13 (IDAEP task)

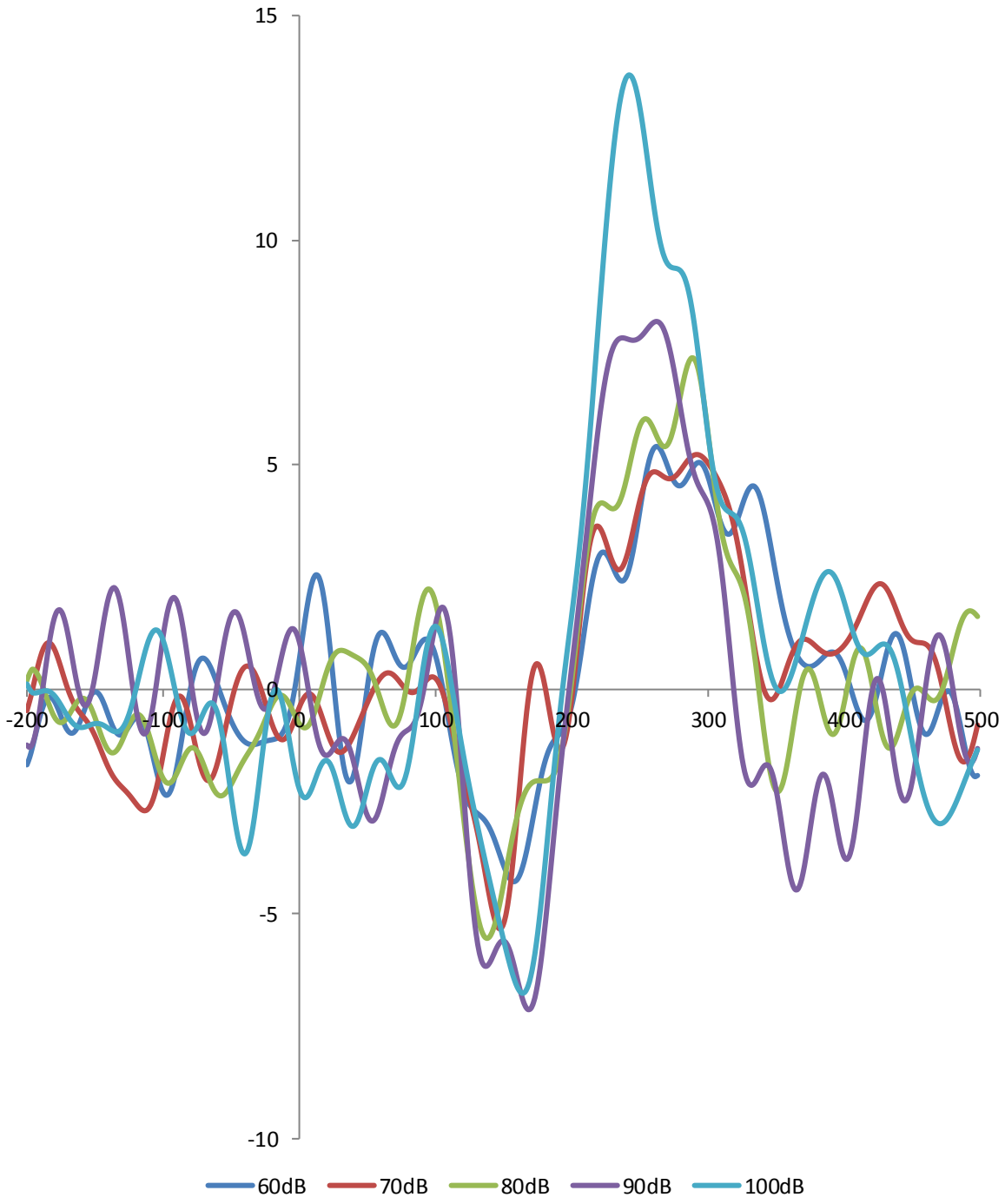

Figure 14. Participant 27 COI back-projected average waveform ExCrit.BV$_{max}$ Config.13 (IDAEP task)

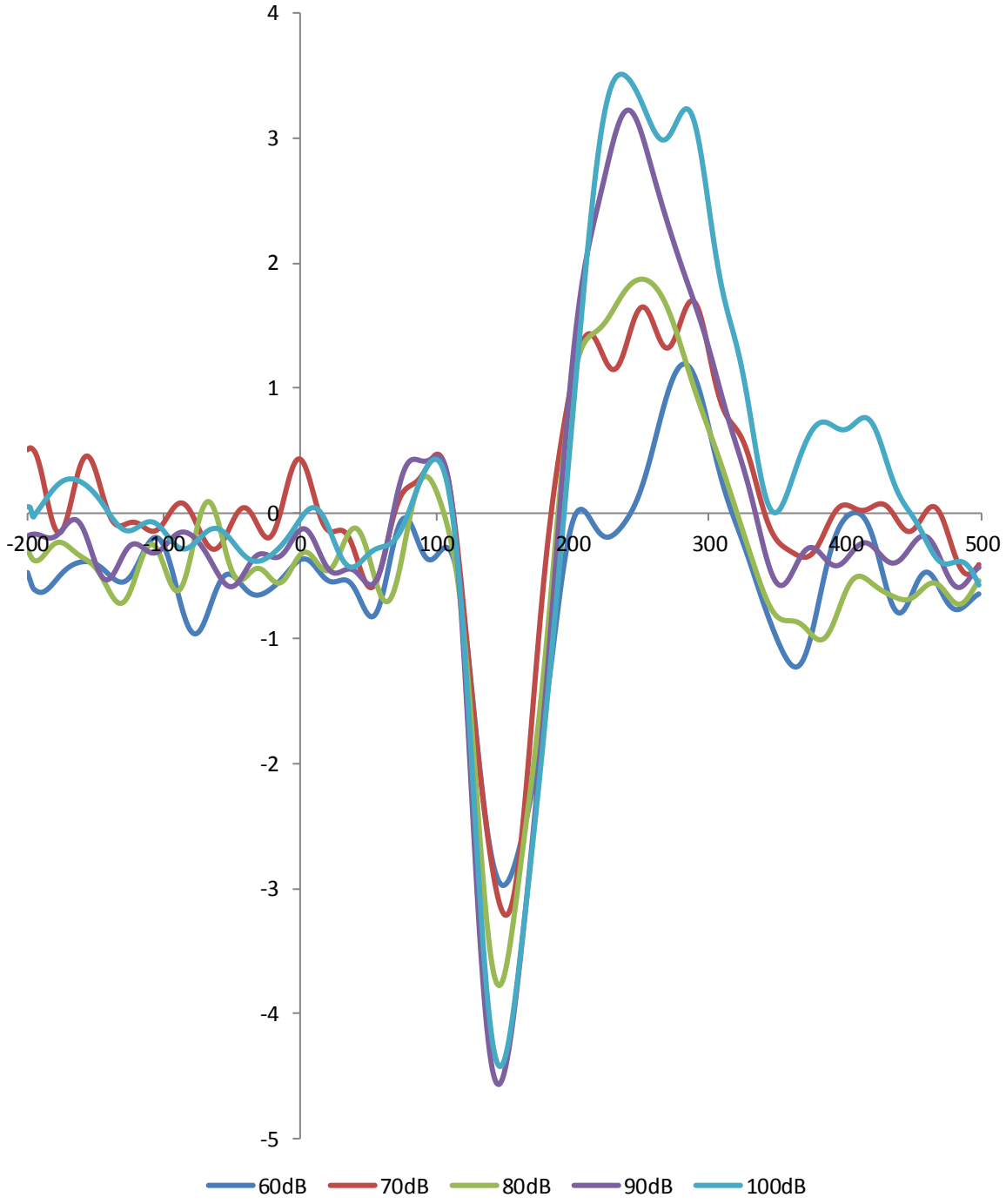

Figure 15. Grand average COI waveform for ExCrit.BV_{max} Config.13 (IDAEP task)

SUMMARY: Based upon this one study, the results suggest that the benefits of data pruning outweigh the potentially negative effects of limiting the total number of epochs entering into the averaging process. It is clear that data preprocessing is not a one size fits all process and different signal analysis techniques require their own preprocessing optimisation. IAPD typically performed the best and was relatively insensitive to minor changes in data preprocessing. ICA-PD performed only slightly better than the wild-data variant of GA-PD, therefore the additional analytical complexity and computational runtime of this ICA-based technique is not considered to be worthwhile. The effectiveness of COIPD varied greatly with configuration and participant selection and was generally a very poor performer. However, under very specific conditions it was observed to outperform all of the other signal analysis techniques and in some instances, may help to uncover otherwise hidden signal characteristics.

Whether the ASF truly follows a polynomial relationship and if P2 splitting is a real feature in the IDAEP at low intensities still requires further investigation. Furthermore, to help clarify some of the results for the different techniques and configurations additional analysis should be performed which utilises participant subsets that are common to all configurations, along with the addition of a further linear modal to directly measure ASF linearity by assessing how the data fits to a model with both a random intercept and a random slope. Firstly, the participant subsets will help to understand the unique nature of the BV_{max} Config.15 dataset and secondly the additional linear model will help to understand how each configuration and technique directly improves signal estimation for each participant rather than assessing how well each participant fits to a common ASF gradient as was performed here.

Chapter 4. Passive Visual Task (LTP)

4.1. Background

Long-term potentiation (LTP) refers to a synaptic mechanism for strengthening neuronal connections following increases in their activity. In 1966, Terje Lømo was the first person to notice how a single 10-20Hz stimulation of the perforant pathway produced a lasting potentiation of excitatory synaptic activity within the dentate granule cells of the hippocampus of anaesthetised rabbits. This research marked the beginning of a long history investigating the neurophysiological underpinnings of LTP and its potential for understanding and explaining the fundamental processes behind information storage within the brain that are necessary for memory and learning. In the years that followed LTP was observed in multiple neocortical areas, specifically motor, somatosensory and visual cortices, where it has been hypothesised to be of functional significance to both motor learning and experience-dependent modification of the visual network during development (Tsumoto and Suda, 1979; Komatsu *et al.*, 1988; Zhang *et al.*, 2000; Heynen and Bear, 2001).

Due to the invasive nature of the in vitro and in vivo techniques typically used to study the mechanisms and phenomenology of LTP in these early animal studies, it wasn't until 1996 that the effect itself was confirmed to be present within the human cortex (Chen *et al.*, 1996). Using isolated human cortical slices, from patients undergoing tissue resection for therapeutic reasons, Chen et al. demonstrated the presence of synaptic plasticity in the inferior and middle temporal cortex. As this region of the brain was already known to be involved in complex visual memory and pattern recognition (Miller *et al.*, 1991; Sakai and Miyashita, 1991; Miyashita, 1993; Miller and Desimone, 1994), the identification of an LTP effect within the same area represented the first true connection between LTP and learning and memory within the human brain. However, as the LTP effect had only ever been elicited by invasive exogenous high frequency stimulation, it wasn't until the effect was induced by more naturalistic stimulation parameters that LTP could be truly argued to be of potential functional relevance. Using a repetitive dimming stimulus presented to the retina of developing tadpoles, Zhang *et al.* (2000) were able to observe LTP within the retinotectal synapses of the visual system, therefore, demonstrating how a real-life exogenous event could induce synaptic plasticity.

The first experiment capable of non-invasively demonstrating LTP in humans was developed by T.J. Teyler and his team in 2005, with variations of this task still being used today (de Gobbi Porto *et al.*, 2015; Smallwood *et al.*, 2015). In the original procedure 128-channel EEG was recorded during repetitive presentation of a visual checkerboard stimulus (~1Hz stimulus onset asynchrony (SOA)) before and after a photic tetanus condition (~10Hz SOA) was presented to one of the two visual hemi-fields. ICA of the grand averaged visual evoked potentials for each condition decomposed the signal into five key components (P100, N1a, N1b, P2 and P3) (Figure 16). Statistical analysis of the impact of the tetanizing stimulus upon the mean amplitude difference of the different components was conducted via a signal analysis methodology similar to that of the ICA-PD described within this thesis. Under this analysis, the N1b mean amplitude was found to be significantly enhanced in the occipital region contralateral to the tetanized hemifield. A subsequent fMRI study by the same group confirmed the N1b components location within the V2 extrastriate area (Clapp *et al.*, 2005), with further studies detailing the specificity of the observed N1b enhancement to the physical characteristics of the potentiating stimulus (McNair *et al.*, 2006; Ross *et al.*, 2008).

Although the functional significance of the N1 component is well understood to represent discriminative processing within the visual cortex (Mangun and Hillyard, 1990; Vogel and Luck, 2000; Hopf *et al.*, 2002), the functional significance of LTP-like potentiation of the N1b component is yet to be fully understood. One potential behavioural outcome that has been noted following N1b potentiation is that there appears to be a slight improvement in the visual detection thresholds for the tetanized stimuli (Clapp *et al.*, 2012). However, given the non-naturalistic approach by which LTP is induced within these non-invasive human studies, the true functional relevance of this finding is somewhat questionable. That being said, there is evidence to suggest that LTP-like effects within the visual cortex of mice can also be generated by techniques that more closely relate to real life visual stimulus exposure rates (Cooke and Bear, 2012). Furthermore, this finding has led some researchers to postulate that the N1b potentiating effect, that has been found during human studies, might truly represent a neural correlate of LTP and that it is therefore possible that this mechanism may form the foundation of perceptual learning within the visual cortex (Clapp *et al.*, 2012).

Having an experimental methodology capable of non-invasively assessing the degree of synaptic plasticity within the visual cortex has led to its use an assessment tool for

general cortical plasticity within patient populations. Using slight variations of the original sensory-induced potentiation task, the potential role of synaptic plasticity dysfunction in the pathophysiology of schizophrenia (Cavus *et al.*, 2012; Mears and Spencer, 2012) and major depression (Normann *et al.*, 2007) have both already been investigated and revealed interesting results. Specifically, in one study of patients with schizophrenia, PCA of the visual evoked potentials (VEPs) to tetanised stimuli demonstrated significant potentiation of the C1 and N1b components in HCs. However, no significant potentiation was observed for either component within the schizophrenic patient group. As schizophrenia is known to be linked with many genetic abnormalities in synapse maturation and neuronal plasticity (Harrison and Weinberger, 2005; Sebat *et al.*, 2009), the ability to detect deficient cortical plasticity in this way highlights the potential clinical usefulness of the LTP task to assess schizophrenic risk and potentially responses to treatment.

The aim of the study presented within this thesis was therefore to investigate the impact of a range of data preprocessing and signal analysis techniques upon the assessment of the LTP effect. Improving the estimation and definition of the underlying potentiated signal will improve the sensitivity of the task along with its potential clinical usefulness.

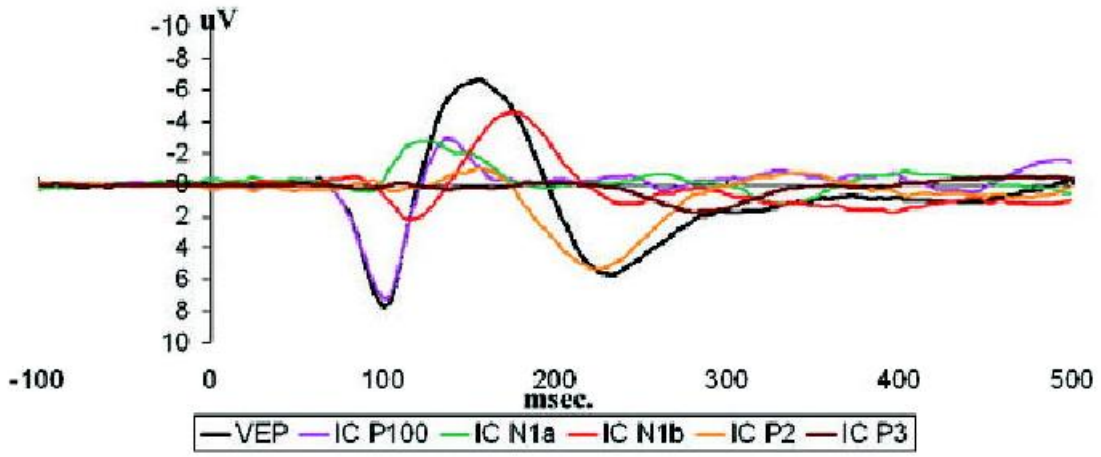

Figure 16. The five ICA components of the visual evoked response, to a checkerboard stimulus, from Teyler *et al.* (2005)

4.2. Methods

4.2.1. *Participants*

30 (18 female) right-handed participants with a mean age of 47.4 (SD 9.1) and normal or corrected to normal vision were recruited as HCs within the ADD study and asked to perform a passive visual LTP task. As for the IDAEP task, participant screening, consent and payment of a small honorarium was completed and conducted as a part of the ADD study.

4.2.2. *Task*

The LTP task consisted of repeatedly presenting a black and white circular checkerboard visual stimulus (diameter 8°, check size 0.3°) to the participant at two different stimulus presentation rates. The task was separated into 5 blocks: 2 pre-tetanus, 1 tetanus and 2 post-tetanus. Pre-tetanus and post-tetanus blocks each consisted of 100 stimuli each presented for 33.3ms and followed by a blank grey screen with a variable ITI of 1000-1500ms. Stimulus presentation therefore occurred at a rate of ~0.8Hz during these test blocks. The tetanus block consisted of 1000 stimuli each presented for 33.3ms with a variable I.T.I of 66.6-100ms, resulting in a presentation rate of ~8.6Hz. Each block lasted approximately 2 minutes and was followed by a rest break of approximately 2 minutes. During the task participants were seated 57cm from the display (resolution 1280x1024; 60Hz refresh rate) in a light-controlled environment and asked to focus on a red fixation square in the centre of the screen.

4.2.3. *EEG Recording*

EEG data was recorded under the same conditions and parameters as set out in the previous chapter for the IDAEP task with evoked potentials being recorded for the pre and post tetanus conditions.

4.3. Analysis

Data preprocessing was performed in accordance with the POP schematic detailed in Figure 2 (page 26) with the soft parameter settings for each configuration as detailed in Table 2. The channel of interest (Oz) for the task was selected as it consistently displayed the largest and clearest visual evoked potential response to the stimulus. Stimulus-locked epochs were made for all trials from -200ms to 830ms. This allowed for the maximum amount of EEG data to be present for training the ocular artefact

algorithms within the gamma POP configuration without any overlapping data between trials. All epochs were subsequently reduced in size (-200ms to 500ms post-stimulus) prior to baseline correction and the assessment and removal of epochs with voltages exceeding +/- 50 μV. As for the previous task the BV, SV and SNR were calculated for all the participant datasets for all of the various POP configurations with participant exclusion via either the $Epoch_{min}$ or BV_{max} exclusion threshold methodologies being applied to the datasets.

Signal analysis focused on the mean amplitude change of the N1b component which was defined as the period of time from the N1 peak to the halfway point between the N1 and P2 peaks (McNair *et al.*, 2006). This approach was used for all four of the signal analysis techniques, even in the case of ICA-based analysis as isolation of a single N1b component was not possible (for more details refer to section 4.5). Statistical analysis focused upon the mean amplitude N1b difference between the 2nd pre-tetanus block (Pre2) and the 1st post-tetanus block (Post1). The statistical significance of N1b potentiation was tested for each all of the different POP and SAP configurations via a paired t-test with the effect size determined by Cohen's d.

4.4. Results

4.4.1. *BV exclusion criteria threshold and PVAF window limits (wild data)*

Of the GA-PD results from the beta-type, wild POP variant data configurations the largest significant effect size was observed with configuration 9 (d=-0.2406 (4dp)) (Table 16). Therefore a 0.5Hz HPF and 30Hz LPF was selected for inclusion within the final gamma-type preprocessing optimization stage along with the two types of ocular artefact correction. Of the two gamma-type configurations, the OAR based upon CRLS regression produced the highest SNR (0.4753 (4dp)), however the LMS regression based OAR of configuration 11 produced the greater effect size (d=-0.2275 (4dp) vs. d=-0.2255 (4dp)). As expected the LTP task elicited the N1 and P2 VEPs and these components were clearly identifiable and maximal within channel Cz of the grand averaged conditional data (Figure 18). Furthermore, the wild-type data demonstrated the characteristic increased N1b negative potentiation for the post-tetanus stimuli (Figure 19).

The median amplitude change of the N1b was calculated for configuration 11 by GA-PD analysis and used as the SOI when defining the BV_{max} exclusion threshold (2.4472 (4dp)). Furthermore, the N1b time window, as derived from the mean N1 (160ms) and P2 (252ms) peak latencies from the four test conditions (Pre1, Pre2, Post1 and Post2) of configuration 11, was used to define the PVAF window limits for the ICA-based signal analysis (PVAF window: Lower Bound 160ms(-20ms), Upper Bound 206ms(+40ms)).

As for the IDAEP task, the relationship between the POP outcome variables (BV, SV and SNR) and the effect size of the wild-type GA-PD signal analysis were tested, however, none of the Pearson's correlations proved to be significant (Figure 17).

Table 16. Wild POP variant preprocessing outcome measures and GA-PD analysis results

POP Configuration						Pre 2 vs Post 1	
Number	Type	N	BV	SV	SNR	p-value	Cohen's d
1	Raw	30	1.6920	52.6473	0.3014	2.61E-04	-0.2791
2	Alpha	30	1.4166	52.3440	0.3071	6.58E-03	-0.1646
3	Beta	30	1.4180	52.2717	0.3420	6.27E-03	-0.1657
4	Beta	30	1.5452	45.0371	0.3271	6.42E-03	-0.1484
5	Beta	30	0.8289	51.7571	0.4111	6.84E-03	-0.1723
6	Beta	30	1.1865	52.0930	0.3477	1.36E-02	-0.1507
7	Beta	30	0.8304	51.6850	0.4743	6.63E-03	-0.1735
8	Beta	30	1.1879	52.0207	0.3926	1.32E-02	-0.1518
9	Beta	30	0.9595	44.4386	0.4739	3.18E-04	-0.2406
10	Beta	30	1.3148	44.7784	0.3813	7.04E-04	-0.2175
11	Gamma	30	0.9502	43.6330	0.4708	5.82E-04	-0.2275
12	Gamma	30	0.9532	43.8623	0.4753	5.61E-04	-0.2255

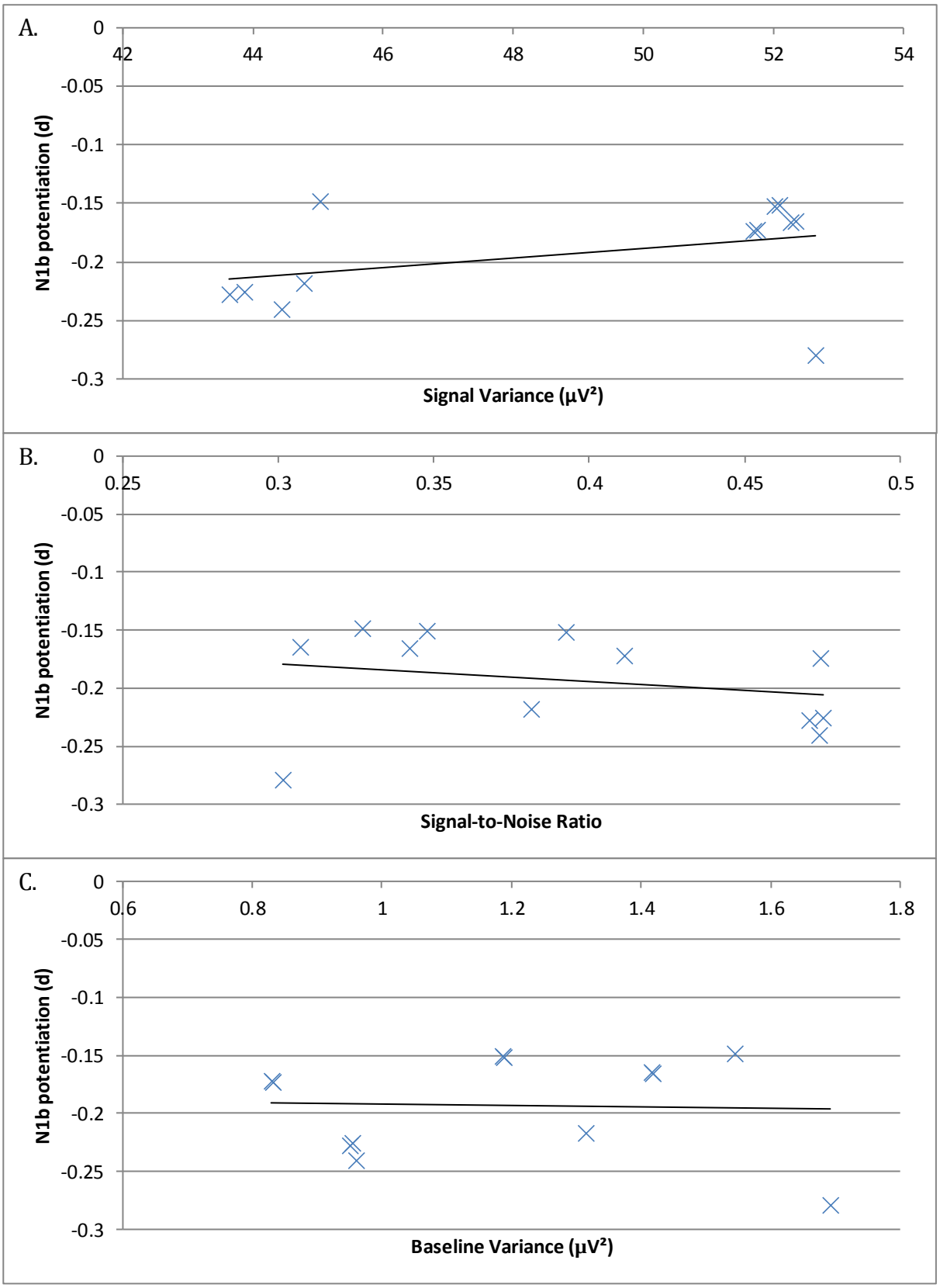

Figure 17. Wild GA-PD POP mean N1b amplitude potentiation effect size (Cohen's d) for each configuration (LTP task) and their relation to A) Signal Variance (r=0.3838. p=0.2133), B) Signal-to-noise ratio (r=-0.2406, p=0.4483) and C) Baseline Variance (r=-0.0409, p=0.8991). Note that no significant correlations were observed between mean N1b amplitude potentiation and any of the POP output measures.

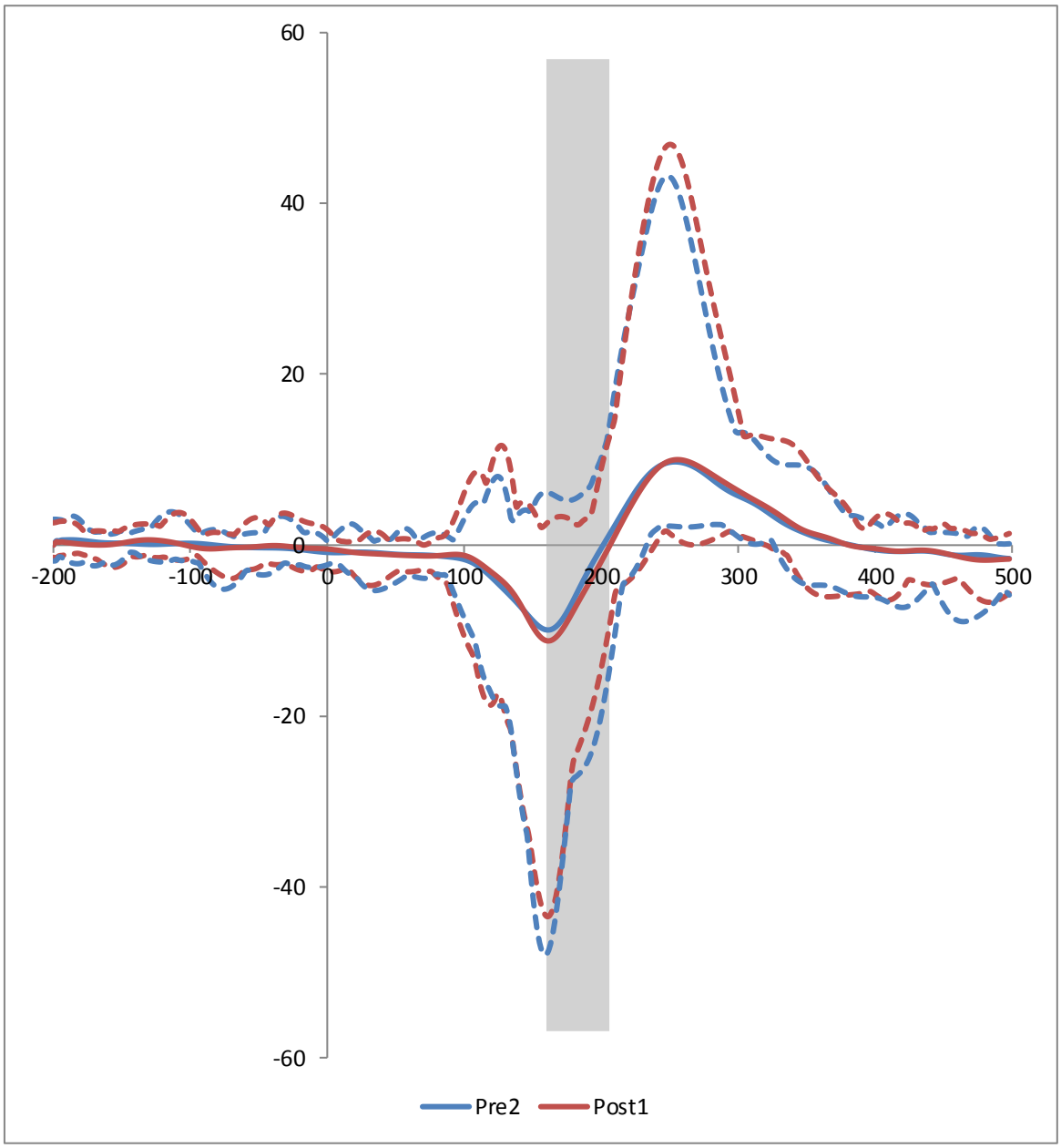

Figure 18. Wild-type grand average ERP waveform for Config.11 (LTP task) Where the ordinate relates to amplitude (μV) and the abscissa relates to time (ms). Dashed lines represent the maximal and minimal individual amplitude values. Grey bar represents the average N1b time window for the two conditions (Pre2 N1b: 160ms to 206ms ; Post1 N1b: 160 to 207 ms).

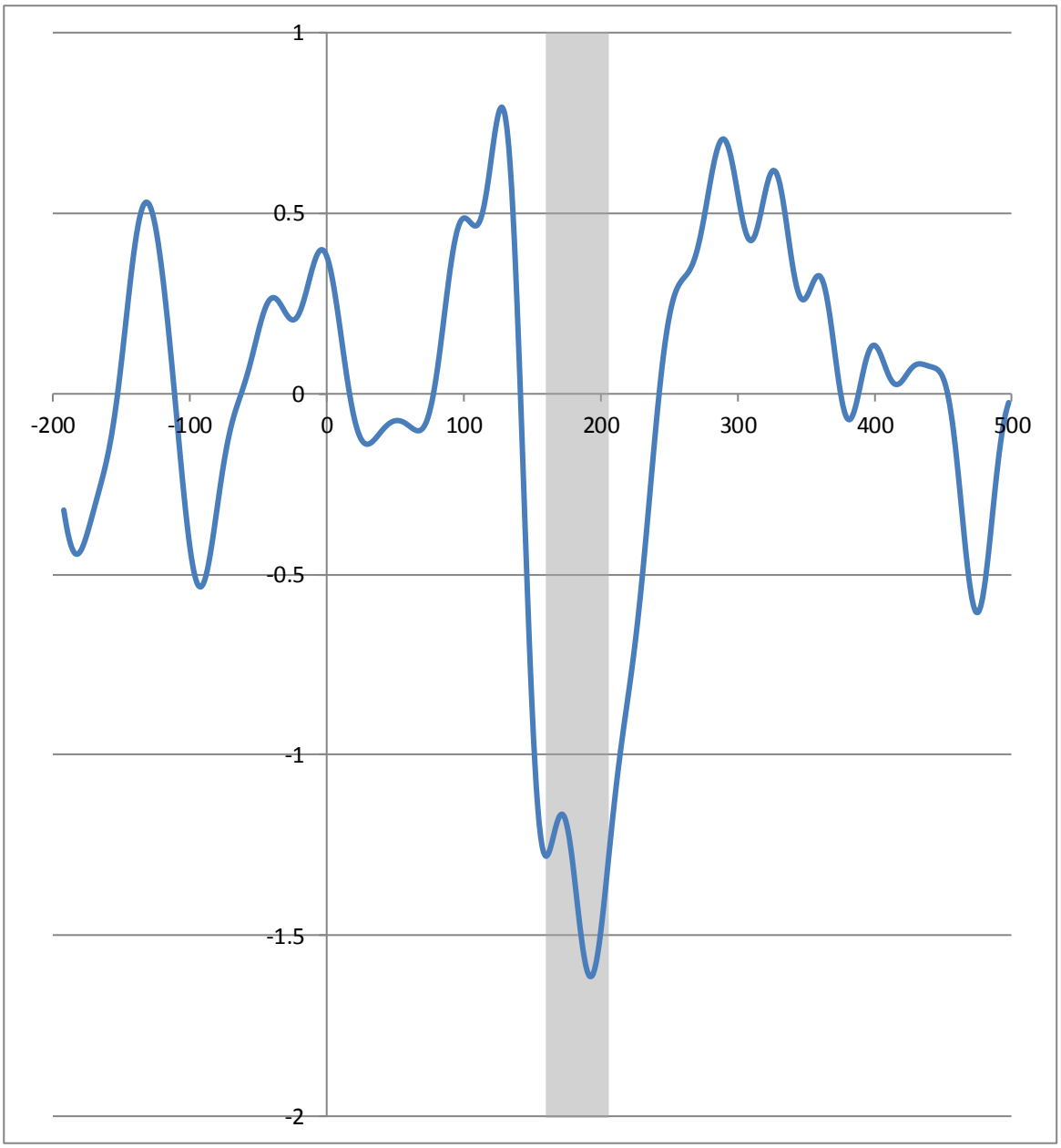

Figure 19. Wild-type Grand average ERP difference waveform (Post1-Pre2) for Config.11. (LTP task) Where the grey bar represents the average N1b time window for the two conditions (Pre2 N1b: 160ms to 206ms; Post1 N1b: 160ms to 207ms). Note the increased negativity of the difference waveform in the timeframe associated with the N1b.

4.4.2. *POP output (pruned data)*

As for the IDAEP task, the analysis proper was performed using data from the pruned variant of the POP with the reporting and interpretation of the results being restricted to those POP configurations where at least half of the participants remained following participant exclusion (i.e. when N>30/2). A total of 16 different POP configurations were tested along with two different methodologies for participant exclusion. Rather than simply selecting one set of filter settings from the beta-type POP configuration to carry forward into the gamma-type preprocessing arrangement, the full range of HPF parameter settings where selected in combination with the best LPF parameter setting (30Hz). These filter settings consistently produced datasets with high SNRs as well as giving the best participant retention rates (Table 17 and Table 18). Furthermore, extending the gamma-type configuration with these extra filter combinations matched the POP analysis performed for the IDAEP task, therefore allowing easier cross-task interpretation of the overall results.

The best BV, SV and SNR results were each observed across a range of different POP configurations, with the only exception being that configuration 7 (0.5Hz HPF, 30Hz LPF) produced the best SNR results for both of the participant exclusion criteria groups ($ExCrit.Epoch_{min}$ SNR=0.5754 (4dp); $ExCrit.BV_{max}$ SNR=0.5989 (4dp)). As can be expected there was a dramatic difference in the BVs between the two exclusion groups, with the best $ExCrit.BV_{max}$ BV (0.5779 (4dp)) being almost two times better than that of the best $ExCrit.Epoch_{min}$ BV (1.0437 (4dp)). Conversely, SV was consistently higher for POP configurations utilising $Epoch_{min}$ based participant exclusion rather than the BV_{max} approach.

Table 17. Pruned POP variant preprocessing outcome measures for ExCrit.Epoch$_{min}$ (LTP task)

POP Configuration		ExCrit.Epoch$_{min}$			
Number	Type	N	BV	SV	SNR
1	Raw	5	1.4500	33.3238	0.3282
2	Alpha	5	1.2831	33.2158	0.3291
3	Beta	5	2.1010	37.9448	0.3798
4	Beta	6	2.1387	31.4917	0.4078
5	Beta	25	1.3690	44.9480	0.5267
6	Beta	16	1.9329	34.4884	0.3166
7	Beta	24	1.4393	44.1898	0.5754
8	Beta	16	1.7388	36.6959	0.3907
9	Beta	24	1.4609	35.7961	0.5426
10	Beta	14	1.8841	31.3510	0.3428
11	Gamma	28	1.0730	32.6443	0.4636
12	Gamma	28	1.0437	32.7332	0.4694
13	Gamma	28	1.0938	39.2665	0.4846
14	Gamma	28	1.1192	39.7309	0.5011
15	Gamma	27	1.2369	40.0155	0.4450
16	Gamma	26	1.0876	41.2410	0.4713

Table 18. Pruned POP variant preprocessing outcome measures for ExCrit.BV$_{max}$ (LTP task)

POP Configuration		ExCrit.BV$_{max}$			
Number	Type	N	BV	SV	SNR
1	Raw	3	1.2906	44.0464	0.4632
2	Alpha	4	1.2581	34.7353	0.3610
3	Beta	3	1.1245	49.8686	0.6128
4	Beta	4	1.0194	41.0002	0.6289
5	Beta	18	0.6363	37.9482	0.5483
6	Beta	10	0.9159	27.7058	0.3528
7	Beta	18	0.6723	37.2525	0.5989
8	Beta	11	0.9793	31.6927	0.4181
9	Beta	15	0.6885	23.2156	0.4815
10	Beta	8	1.0577	24.9932	0.3632
11	Gamma	22	0.6840	30.2918	0.4995
12	Gamma	23	0.7119	31.6754	0.5148
13	Gamma	22	0.6370	38.1311	0.5351
14	Gamma	22	0.5779	35.8613	0.5301
15	Gamma	21	0.6776	35.7221	0.4594
16	Gamma	21	0.6389	34.6124	0.4595

4.4.3. *GA-PD (pruned data)*

Under $Epoch_{min}$ participant exclusion nine of the sixteen configurations demonstrated a significant N1b amplitude potentiation (Table 19). Of these results the best gamma-type effect size was observed for configuration 15 (No additional HPF, 30Hz LFP and OAR by LMS regression) (d=-0.2501 (4dp)). However, a greater effect size was observed when ocular artefact removal was absent from preprocessing altogether, i.e. during the beta-type POP configuration (Config.7: d=-0.2856 (4dp)).

GA-PD analysis of the BV_{max} participant exclusion group data only generated significant N1b potentiation with configurations 11 and 16 and marginal significance with configurations 5, 7, 13 and 15. However, despite the fact that the BV_{max} methodology led to fewer significant results, the largest effect size for the whole GA-PD technique was actually observed for ExCrit.BV_{max} Config.16 (d=-0.3919 (4dp)) (Figure 20 and Figure 21).

Table 19. Pruned GA-PD N1b potentiation statistics (LTP task). Significant paired t-test results are highlighted in red text. Largest significant effects sizes for each POP configuration are highlighted by red boxes.

POP Configuration		ExCrit.$Epoch_{min}$			ExCrit.BV_{max}		
Number	Type	N	p-value	Cohen's d	N	p-value	Cohen's d
1	Raw	5	5.10E-01	0.0883	3	7.77E-01	0.0737
2	Alpha	5	5.77E-01	0.0815	4	1.20E-02	-0.6352
3	Beta	5	7.65E-01	-0.0410	3	2.69E-02	-0.5997
4	Beta	6	5.05E-02	-0.3171	4	3.38E-01	-0.1478
5	Beta	25	4.69E-02	-0.1847	18	5.80E-02	-0.2836
6	Beta	16	2.59E-01	-0.1402	10	4.63E-01	-0.1940
7	Beta	24	2.58E-03	-0.2856	18	7.55E-02	-0.2297
8	Beta	16	5.42E-01	-0.0705	11	1.03E-01	-0.4164
9	Beta	24	1.35E-02	-0.2203	15	1.16E-01	-0.2429
10	Beta	14	5.32E-01	-0.0556	8	3.92E-03	-0.8136
11	Gamma	28	3.98E-02	-0.1626	22	6.05E-03	-0.3237
12	Gamma	28	4.84E-02	-0.1490	23	1.99E-01	-0.1364
13	Gamma	28	5.97E-03	-0.2304	22	8.39E-02	-0.2197
14	Gamma	28	1.13E-02	-0.2140	22	1.31E-01	-0.1792
15	Gamma	27	8.72E-03	-0.2501	21	6.75E-02	-0.2587
16	Gamma	26	2.53E-02	-0.2013	21	4.29E-03	-0.3919

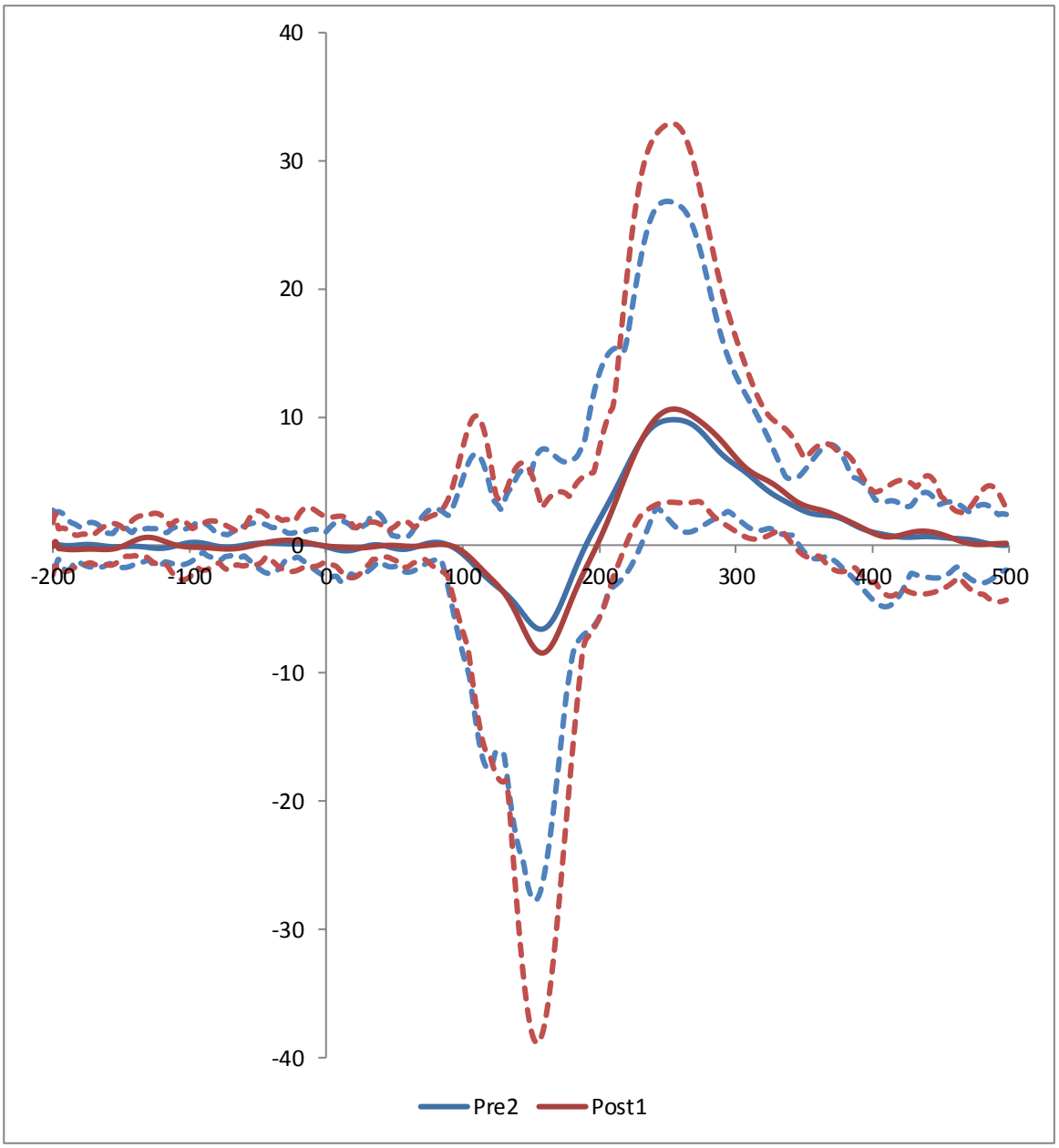

Figure 20. Grand average ERP waveforms for ExCrit.BV_{max} Config.16 (LTP task). Where the ordinate relates to amplitude (μV) and the abscissa relates to time (ms). Dashed lines represent the maximal and minimal individual amplitude values.

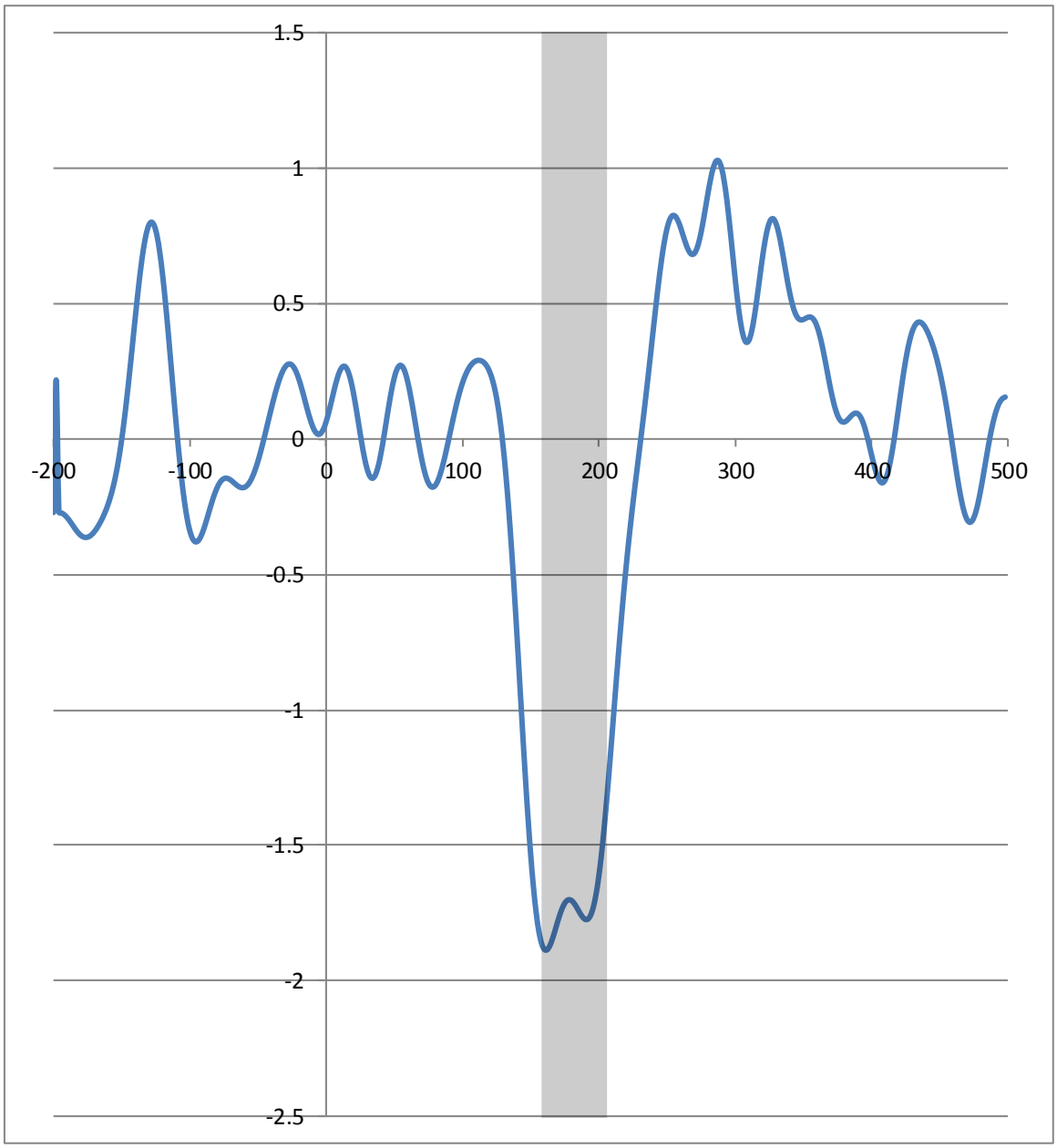

Figure 21. Grand average ERP difference waveform (Post1-Pre2) for ExCrit.BV$_{max}$ Config.16 (LTP task). Where the grey bar represents the average N1b time window for the two conditions (Pre2 N1b: 158ms to 206ms; Post1 N1b: 158ms to 206ms).

4.4.4. *IAPD*

Signal analysis of the $Epoch_{min}$ exclusion group datasets by IAPD produced four significant N1b potentiation results and 1 of marginal significance (Table 20). BV_{max} threshold exclusion again resulted in fewer significant results and more potentiation trends (Configs. 11 and 16 $p<0.05$; Configs. 14 and 15 $p<0.1$). The largest effect size in both exclusion criteria groups was observed for configuration 16 ($ExCrit.Epoch_{min}$ d=-0.1736; $ExCrit.BV_{max}$ d=-0.2309). However, in comparison with the best GA-PD results, IAPD appeared to diminish the strength of the potentiation effect (Δ Effect Size: $ExCrit.Epoch_{min}$ -39.22% (2dp), $ExCrit.BV_{max}$ -41.09% (2dp)).

Table 20. IAPD N1b potentiation statistics (LTP task). Significant paired t-test results are highlighted in red text. Largest significant effects sizes for each gamma type POP configuration are highlighted by red boxes.

POP Configuration		$ExCrit.Epoch_{min}$			$ExCrit.BV_{max}$		
Number	Type	N	p-value	Cohen's d	N	p-value	Cohen's d
1	Raw	5	6.01E-01	-0.1205	3	9.72E-01	-0.0176
2	Alpha	5	9.72E-01	0.0071	4	8.92E-01	-0.0489
3	Beta	5	7.37E-01	0.0558	3	7.67E-01	-0.1822
4	Beta	6	7.37E-01	-0.0701	4	6.15E-01	-0.2540
5	Beta	25	1.47E-01	-0.1139	18	1.22E-01	-0.1883
6	Beta	16	3.49E-01	-0.1006	10	5.98E-01	-0.0997
7	Beta	24	2.59E-02	-0.1541	18	1.46E-01	-0.1538
8	Beta	16	8.45E-01	0.0393	11	7.58E-02	-0.2667
9	Beta	24	4.96E-01	-0.0868	15	2.85E-01	-0.1517
10	Beta	14	3.02E-01	-0.0954	8	2.36E-01	-0.2313
11	Gamma	28	1.11E-01	-0.1088	22	3.85E-02	-0.1856
12	Gamma	28	6.81E-01	-0.0396	23	1.69E-01	-0.1262
13	Gamma	28	4.98E-02	-0.1393	22	1.21E-01	-0.1520
14	Gamma	28	4.92E-02	-0.1577	22	8.99E-02	-0.1948
15	Gamma	27	6.04E-02	-0.1586	21	9.88E-02	-0.2033
16	Gamma	26	2.56E-02	-0.1736	21	3.58E-02	-0.2309

4.4.5. *ICA-PD*

Similar to the ICA analysis reported for the IDAEP study, only a limited subset of the POP configurations where selected for analysis via ICA-PD and COIPD. The five configurations (Configs. 7, 11, 13, 15 and 16) were selected based upon the strength of their performance during the GA-PD and IAPD analysis stages. Out of these five configurations two of the ICA-PD $ExCrit.Epoch_{min}$ results were significant and three of the ExCrit.BVmax results were significant (Table 21). For both of the participant exclusion groups the largest significant effect size was elicited from the sole beta-type POP configuration selected for analysis by ICA ($ExCrit.Epoch_{min}$ d=-0.1790; $ExCrit.BV_{max}$ d=-0.2258). However, again neither of these ICA-PD approaches produced an effect size larger than that which was obtained via the GA-PD signal analysis procedure (Δ Effect Size: $ExCrit.Epoch_{min}$ -37.32% (2dp), $ExCrit.BV_{max}$ -42.38% (2dp)).

Table 21. ICA-PD N1b potentiation statistics (LTP task). Significant paired t-test results are highlighted in red text. Largest significant effects sizes for the beta and gamma configurations of the ICA analysis subset are highlighted by red boxes.

POP Configuration		$ExCrit.Epoch_{min}$			$ExCrit.BV_{max}$		
Number	Type	N	p-value	Cohen's d	N	p-value	Cohen's d
7	Beta	24	2.55E-02	-0.1790	18	8.31E-03	-0.2258
11	Gamma	28	9.95E-02	-0.1219	22	2.19E-02	-0.2025
13	Gamma	28	3.58E-02	-0.1593	22	4.88E-02	-0.2013
15	Gamma	27	5.25E-02	-0.1295	21	1.28E-01	-0.1363
16	Gamma	26	1.24E-01	-0.1155	21	6.98E-02	-0.1769

4.4.6. *COIPD*

For both of the participant exclusion groups COIPD resulted in significant N1b potentiation for the datasets preprocessed with the same three POP configurations (Configs. 7, 11 and 13) (Table 22). As for the ICA-PD analysis the largest significant effect sizes were observed for configuration 7 ($ExCrit.Epoch_{min}$ d=-0.1819; $ExCrit.BV_{max}$ d=-0.2469) (Figure 22 and Figure 23) and again the best results from this technique were weaker than the best results obtained via GA-PD (Δ Effect Size: $ExCrit.Epoch_{min}$ -36.31% (2dp), $ExCrit.BV_{max}$ -37.00% (2dp)).

Table 22. COIPD N1b potentiation statistics (LTP task). Significant paired t-test results are highlighted in red text. Largest significant effects sizes for the beta and gamma configurations of the ICA analysis subset are highlighted by red boxes.

POP Configuration		$ExCrit.Epoch_{min}$			$ExCrit.BV_{max}$		
Number	Type	N	p-value	Cohen's d	N	p-value	Cohen's d
7	Beta	24	3.20E-02	-0.1819	18	1.91E-02	-0.2469
11	Gamma	28	4.07E-02	-0.1437	22	1.70E-02	-0.2281
13	Gamma	28	1.29E-02	-0.1723	22	9.15E-03	-0.2394
15	Gamma	27	1.01E-01	-0.0856	21	2.63E-01	-0.0655
16	Gamma	26	1.92E-01	-0.0832	21	1.60E-01	-0.1163

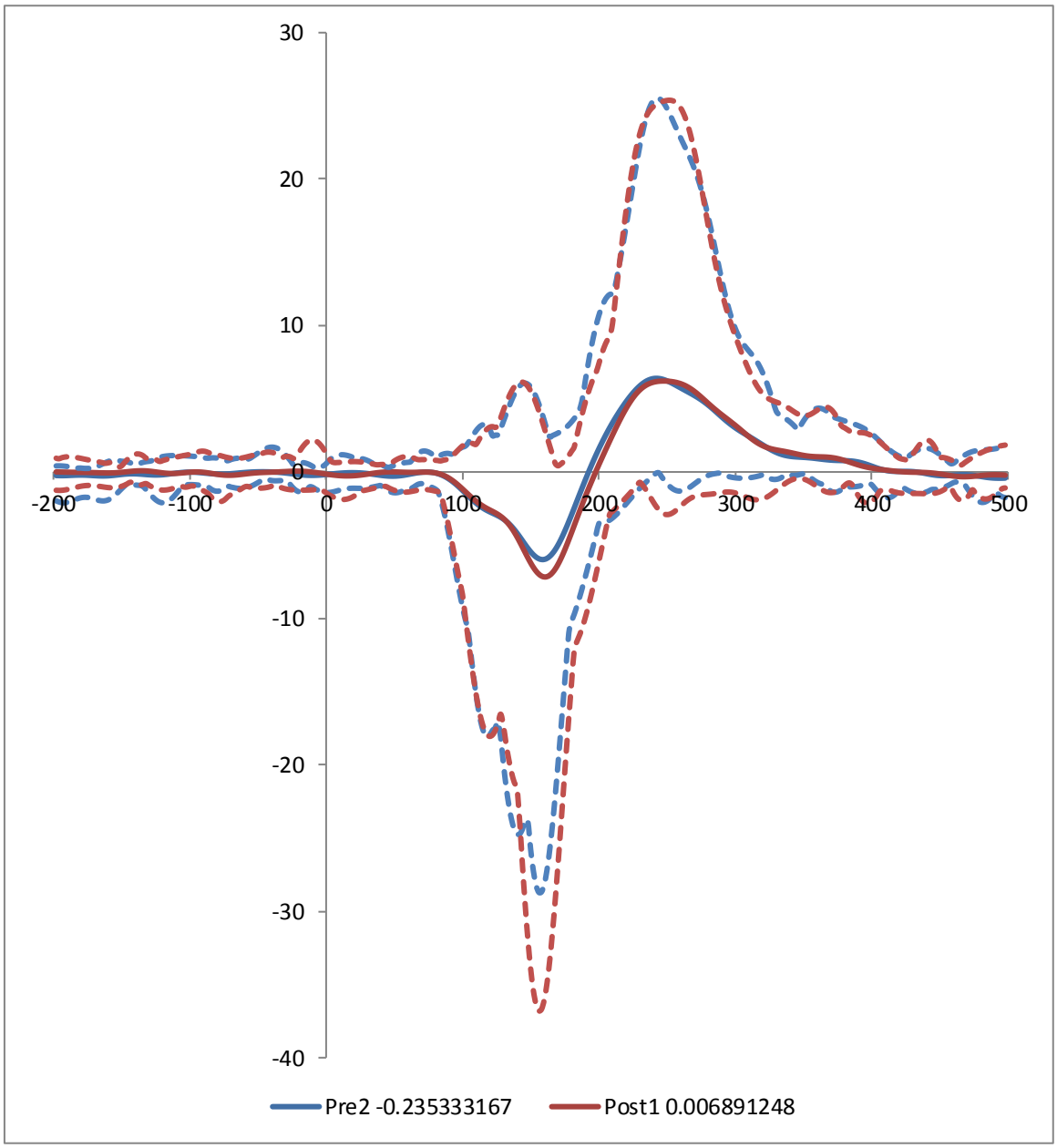

Figure 22. Grand average back-projected COI waveforms for ExCrit.BV_{max} Config.7 (LTP task). Where the ordinate relates to amplitude (μV) and the abscissa relates to time (ms). Dashed lines represent the maximal and minimal individual amplitude values.

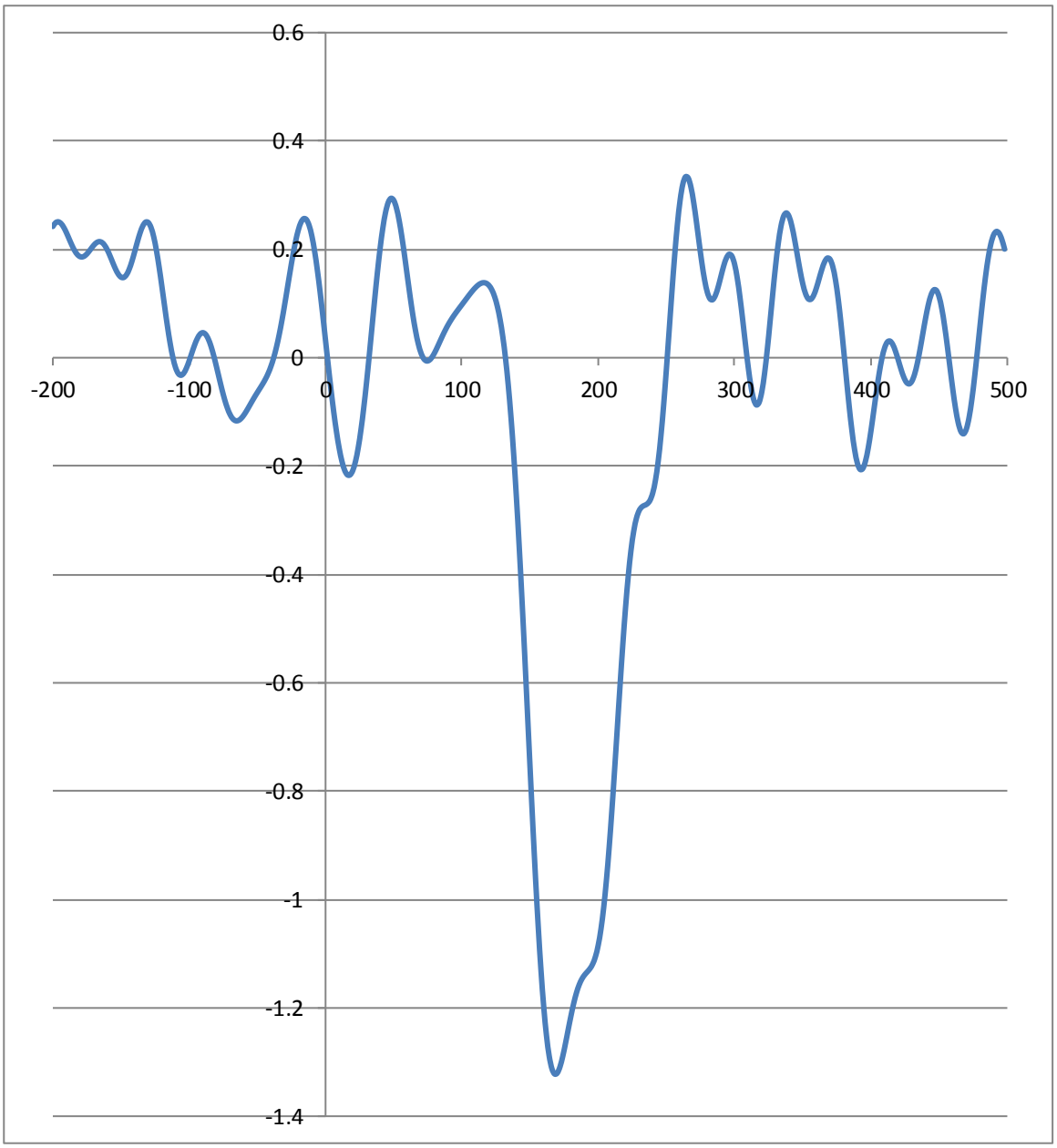

Figure 23. Grand average back-projected COI difference waveform (Post1-Pre2) for ExCrit.BV_{max} Config.7 (LTP task)

4.5. Conclusions

It is clear from the wide range of effect sizes observed for the many different POP and SAP analysis configurations that the isolation and estimation of N1b potentiation for the LTP task can be substantially affected by the specific choice of preprocessing and signal analysis approach (Table 23). In general, the potentiation of the N1b component elicited by the task only demonstrated a small-medium effect size across all of the tested analysis configurations (Cohen, 1988; Cohen, 1992). Furthermore, the preprocessing outcome measures (BV, SV and SNR) for these configurations did not appear to be able to predict the best parameter settings for obtaining the best effect size within the wild-type GA-PD datasets (Figure 17). However, rather than being due to any potential inadequacies in the ability of the outcome measures to assess the quality of the data, the lack of a significant correlation with the task-specific effect sizes may in fact highlight potential inadequacies in the GA-PD approach as well as unique artefact contributions within the wild-type dataset.

Table 23. POP configuration subset effect sizes (Cohen's d) for each signal analysis technique (LTP task). Colour coding relates to effectiveness of the technique for a given dataset from best to worst (green > yellow > orange > red)

	Config.	N	GA-PD	IAPD	ICA-PD	COIPD
$Epoch_{min}$	7	24	-0.2856	-0.1541	-0.1790	-0.1819
	11	28	-0.1626	Not sig.	-0.1219 (trend)	-0.1437
	13	28	-0.2304	-0.1393	-0.1594	-0.1723
	15	27	-0.2501	-0.1586 (trend)	-0.1295 (trend)	Not sig.
	16	26	-0.2013	-0.1736	Not sig.	Not sig.
BV_{max}	7	18	-0.2297 (trend)	Not sig.	-0.2258	-0.2469
	11	22	-0.3237	-0.1856	-0.2025	-0.2281
	13	22	-0.2197 (trend)	Not sig.	-0.2013	-0.2394
	15	21	-0.2587 (trend)	-0.2033 (trend)	Not sig.	Not sig.
	16	21	-0.3919	-0.2309	-0.1769 (trend)	Not sig.

For instance, the largest effect size within the wild-type GA-PD analysed data was observed for the raw EEG dataset (d=-0.2791), however, when the same data was subsequently preprocessed to remove electrical line noise the effect size of the task was greatly reduced (d=-0.1646). As the N1b component and electrical line noise relate to activity within very different spectral frequencies, the reduction in the effect size of the N1b following line noise removal suggests that the effect size measure was somehow being positively biased by the electrical line noise.

Furthermore, of the four different signal analysis approaches the GA-PD technique could typically be seen to produce significant potentiation effects with larger effect sizes for a greater number of the different preprocessing configurations. The greatest potentiation effect was observed when the LTP data was first preprocessed with the *cleanline* function, then offline filtered with only a 30 Hz LPF and OAR performed by a CRLS-based regression technique, followed by participant exclusion based upon the BV_{max} exclusion threshold methodology (i.e. POP configuration 16) (Figure 24). However, the fact that the GA-PD technique appeared to specifically outperform the IAPD approach suggests that the generalized GA-PD N1b time window was being positively influenced by neighbouring ERP components in individuals with peak latencies that greatly differ from those of the grand average. For example, if the N1 VEP component was also being potentiated by the tetanizing stimulus then the slightly broader and less specific GA-PD N1b time window could have easily been biased by any enhanced N1 activity. It could therefore be argued that signal analysis that uses broad and generalised time frames for mean peak amplitude calculations, such as the GA-PD approach, should be avoided when components of interest are known to closely neighbour (both temporally and spatially) other active ERP components. Additionally, these results from the GA-PD approach highlight the caution with which large effect sizes should be interpreted, as comparatively larger effect sizes do not always appear to relate to improvements in signal isolation or estimation.

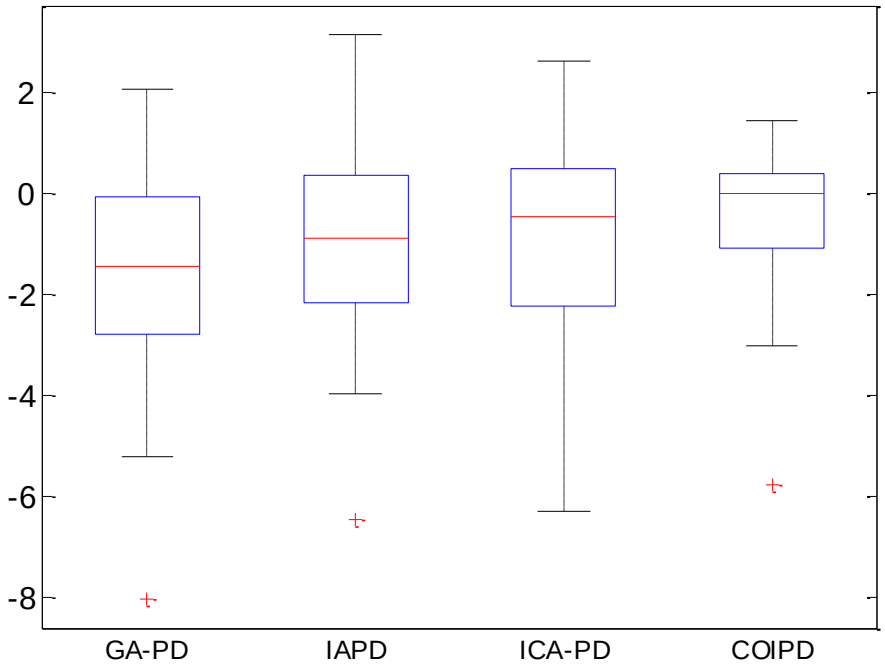

Figure 24. N1b potentiation boxplots for ExCrit.BV$_{max}$ Config.16 (LTP task). Where the ordinate relates to amplitude difference (Post1-Pre2) (µV). Outliers (red marker) correspond to data that was +/-2.7 standard deviations from the mean

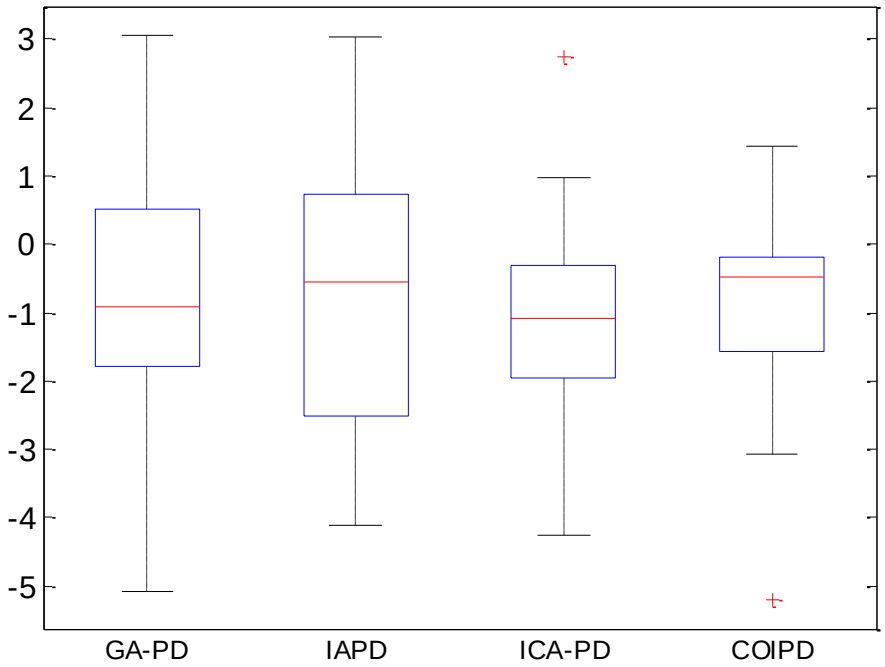

Figure 25. N1b potentiation boxplots for ExCrit.BV$_{max}$ Config.7 (LTP task). For details see Figure 24.

As the GA-PD results likely did not accurately represent the activity of just the N1b component the effectiveness of the remaining techniques have been considered and compared separately. Therefore, with the results of the GA-PD technique set aside the effectiveness of the three other signal analysis techniques largely followed the expected pattern (N1b Effect Size: COIPD > ICA-PD > IAPD) (Table 23). Interestingly the largest N1b potentiation effect size was observed for COIPD signal analysis when the LTP task data was preprocessed with a configuration where ocular artefact reduction was absent from the pipeline (i.e. beta-type POP configuration 7) (Figure 25). The fact that COIPD performed better without OAR (Config. 7 vs Config.13, Table 23) potentially demonstrates the strength of ICA for artefact classification/isolation. However, it may also highlight a potential weakness within the regression based OAR technique, as the results indicate that the OAR process must also be resulting in a small degree of signal loss. Although this would suggest that any EEG task analysis using COIPD could potentially forego various stages of data preprocessing and allow ICA to handle the greater part of artefact isolation/removal this consideration would require the removal or relocation of the pruning stage within the POP. Under the current pipeline design any trials containing large ocular artefacts would be excluded during the data pruning stage and therefore would be prematurely removed from the dataset before the ICA stage could be used to clean the data.

Despite the fact that the data quality indexes weren't observed to significantly correlate with task outcomes for the wild-type GAPD data, visual assessment of the SNRs of the POP subsets for the IAPD, ICA-PD and COIPD data outcomes appeared to show some positive correlation between effect sizes and SNRs of the best performing datasets. Furthermore, the BV measure may also show promise in relation to its use within the BV_{max} participant exclusion approach. As it was the analysis strategies that incorporated the BV_{max} participant exclusion approach that typically displayed the largest significant effect sizes.

Despite both of the ICA based analysis techniques successfully demonstrating LTP within a time frame relating to the N1b, the fact that a separate N1b component was not observed in any of the ERP signal decompositions needs to be addressed. In the original visual LTP task by Teyler *et al.* (2005), the VEP was separated by ICA into 5 key components, with the N1b being one of them. One fundamental difference between the visual LTP task presented within this thesis and the one used by Teyler's group is that

the later used a hemi-field visual checkerboard rather than a full circular checkerboard stimulus. As previous research has shown the N1b component to be generated bilaterally within the extra striate visual cortex (Clapp *et al.*, 2005), during hemi-field visual presentation (Teyler *et al.*, 2005), it therefore has a very unique spatial topography when compared to the evoked N1 and P2 components which predominately occur within primary visual cortex contralateral to the visual stimulus (Di Russo *et al.*, 2002; Luck, 2014). The present version of the task therefore doesn't have the ability to produce these unique topographies as the centrally presented full visual checkerboard stimulus leads to the N1, P2 and N1b components to all be generated bilaterally. As a result the unique spatial topography and activity of the N1b component cannot be disentangled from the N1 and P2 components and this is likely why only a general VEP ICA component was found within the occipital region for this data. Furthermore, the high-density EEG setup used by Teyler's group would have allowed greater source localisation and separation due to the greater amount of spatial information available for the task.

SUMMARY: The data from this study has shown that the BV_{max} participant exclusion methodology developed for this thesis has the ability to greatly improve the outcome statistics for a given task regardless of the preprocessing or signal analysis methodology used. The potential negative impact of the GA-PD technique's overgeneralised approach to signal analysis has been observed and demonstrates the caution with which results from this technique should be interpreted. The COIPD signal analysis approach stood out as the best approach for N1b analysis. However, the lack of an individual N1b component during signal decomposition highlights how the use of ICA-based analysis techniques should always be preceded by careful consideration of the specific stimulus features that make up the task to be able to fully leverage their potential.

Chapter 5. Discussion

Within this thesis event-related EEG analysis was assessed using real world data from two sensory evoked potential studies. The investigation of different EEG analysis procedures was undertaken through the consideration and configuration of two semi-automated EEG analysis pipelines for data preprocessing and signal isolation/estimation (i.e. the POP and SAP). Different combinations of preprocessing techniques and participant exclusion methodologies were performed upon the EEG data for each task to try and ascertain the optimal preprocessing strategy. Four different signal analysis methodologies (GA-PD, IAPD, ICA-PD and COIPD) were then performed upon a subset of the preprocessed datasets for each task with task-specific outcome statistics being used to measure the success of each technique.

The primary aim of this thesis was then to compare the effectiveness of the relatively complex ICA-based signal analysis strategies against the simpler event-related trial-averaging ERP based approaches. A secondary aim was to assess the sensitivity of these techniques to the different data preprocessing strategies. Furthermore, with the introduction of the novel BV-based participant exclusion criteria, the effectiveness of this approach in comparison to the standard minimum epoch number threshold technique was also of interest.

5.1. Conclusions

5.1.1. *Signal analysis strategies*

In considering all of the results from both tasks it is evident that using an ICA-based signal analysis approach does not inherently guarantee that the event-related neuronal signals for a given task will be estimated with any greater accuracy than could have been achieved with a simpler trial-averaging ERP waveform based peak detection approach. The prime example of this being when the COIPD and ICA-PD signal analysis methods both demonstrated lower effect sizes following tests of the IDAEP ASF linearity for the majority of preprocessed datasets. If the ASF of the underlying neuronal activity is truly a linear relationship, then the observed decrease in linear effect size for the ICA-based approaches signifies that these techniques have failed to decrease the amount of variance in the signal estimate due to either a poor isolation or classification of the

neuronal source activity that generates the intensity dependent auditory evoked responses.

P2 peak classification within the IDAEP task was noted for being just as difficult and highly subjective for the back-projected ICA component activity for each participant as it was for peak detection of the individual conditional ERP average waveforms. This observation doesn't support the expected benefits of using BSS techniques to better isolate event-related signals. Therefore, it could be argued that ICAs inability to better isolate the auditory evoked potentials or specifically resolve the P2 peak splitting phenomenon that appeared to be present for low-intensity stimuli may be an indication that the algorithm didn't have enough data to accurately decompose the underlying signals. As the EEG data for both tasks was only recorded using a 32 channel EEG setup the ICA algorithm would have been both limited in the number of independent sources that it could decompose the data into in addition to there also being less topographical information present to identify unique source distributions within. Had the tasks been performed using a high-density EEG recording setup (i.e. 128 channels or more) the ICA-based approaches may have been able to isolate the N1 and P2 auditory evoked potentials with greater accuracy as well as potentially explain the source of the P2 ERP component ambiguity.

For the LTP task both of the ICA-based signal analysis approaches typically produced larger task-specific effect sizes than when the N1b component was assessed by IAPD. Furthermore, the fact that the N1b wasn't isolated by ICA into an independent component that was separate from the N1 and P2 primary visual evoked responses suggests that the full potential of the ICA-based techniques wasn't realized under the current experimental conditions. For instance, the ICA-based techniques could have proved to be more effective had the LTP task been designed to elicit the unique independent source distribution properties of the potentiated visual evoked response. Specifically, previous research has shown that it is possible to separately isolate the N1b component from the N1 and P2 components when the LTP task uses a hemi-field visual stimulus (Teyler *et al.*, 2005). When a hemi-field visual stimulus is presented to a participant the N1 and P2 visual evoked responses are commonly observed to be the most prominent within the primary visual cortex contralateral to the visual stimulus (Clapp *et al.*, 2005; Teyler *et al.*, 2005). Whereas the neuronal activity that is produced when a tetanised hemi-field visual stimulus is presented to a participant is believed to

be generated within the extrastriate visual cortex and can be seen to propagate both contra and ipsi-lateral to the stimulus (Clapp *et al.*, 2005). This clear distinction in the source distributions of the different visual evoked responses (N1/P2 vs. N1b) is likely to be the key element that has enabled the N1b activity to be isolated by previous studies. In this regard, when a full-field tetanised visual stimulus is presented as was performed within this thesis, the primary visual evoked response will be generated in both hemispheres and therefore will become harder to differentiate from the neighbouring bi-laterally generated extrastriate visual cortex activity. This therefore demonstrates the important role that task design can play in ensuring BSS-techniques are leveraged to find unique independent source distribution properties.

The large inflated effect size results that were observed for the LTP task when signal analysis was performed by the GA-PD approach demonstrates how analysis techniques that assess amplitude changes by generalised mean amplitude windowing methodologies can lead to a biased estimation of the SOI and therefore should always be interpreted with caution. The simplicity of the GAPD technique and its overgeneralised approach to defining the N1b time window for all individuals doesn't account for inter-individual differences in peak latency. This approach appears to have allowed the N1 ERP component to occasionally be included within the N1b amplitude estimate leading to an inflated and biased effect size. It is therefore recommended that GAPD analysis should only be used in instances where the ERP component of interest is believed to be well separated from neighbouring components or when there is likely to be only a small amount of inter-individual difference in peak latency for the SOI.

5.1.2. *Preprocessing strategy optimisation and sensitivity*

When considering how the different signal analysis approaches might have impacted upon the magnitude and significance of the various task-specific effect sizes, a clear difference was observed in the range of results that were produced by the two different tasks. For the LTP task data, the overall significance of the task-specific outcome measure was highly dependent upon the preprocessing strategy that was used before signal analysis. However, for the IDAEP task data the different preprocessing strategies only appeared to alter the magnitude of the task-specific results by a fairly small margin. Furthermore, for the IDAEP task the many different preprocessing parameter configurations and participant rejection methodologies, didn't appear to have any great impact upon the significance level of the ASF linear relationship.

In comparing all of the IAPD effect sizes from the gamma type POP configurations, all of the different combinations of HPF cutoffs and OAR methodologies only improved the r^2 effect size of the ASF linear relationship at most by less than 0.04. However, when using the COIPD approach a broader range of effect sizes were elicited by the different HPF cutoffs and participant exclusion methodologies with a maximum improvement of approximately 0.23. This suggests that preprocessing optimization may be more important for accurate BSS and subsequent signal analysis and less of a concern for ERP waveform based approaches. However, this is not to say that preprocessing can be completely ignored when using ERP waveform based approaches. The large number of participants that failed to meet the inclusion criteria following certain preprocessing strategies clearly demonstrated that both low pass filtering at around 30Hz as well as some form of OAR were both valuable preprocessing techniques for ensuring a good amount of quality data existed for the majority of participants. However, as the effectiveness of the *cleanline* function was not directly assessed, the perceived success of the 30Hz LPF in comparison to the 60Hz LPF may in fact have been driven by its ability to remove any 50Hz electrical line noise that may not have been adequately removed by the *cleanline* function. Therefore, without being able to formally assess the effectiveness of the *cleanline* function the specific source of the artefactual activity that caused this preprocessing strategy to perform poorly, can only be speculated upon. For instance, the quality of the data for the preprocessing strategies incorporating the 60Hz LPF, may be due to the presence of poorly removed electrical line noise artefact or if it be the result of the filter letting through other artefact signals and non-event-related neuronal activity within the 30-60hz frequency range.

The use of specific preprocessing parameters, the choice of participant exclusion methodology and the type of signal analysis approach can all be seen to impact upon the perceived success of the LTP task to elicit N1b potentiation. Of the many gamma type POP configurations, approximately half of the statistical tests for the assessment of the N1b potentiation failed to reach significance and the ICA-based approaches occasionally failed to find significant results for datasets were IAPD analysis had been successful (i.e. Configuration 16; No digital HPF, 30Hz LPF, OAR by CRLS). When considering the significant results for all of the different signal analysis approaches (without consideration of the potentially biased GAPD data) the maximum increase in effect size (Cohens d) that was achieved across all of the different datasets was relatively small (~0.1).

The relationship between data quality, as measured by BV, SV and SNR, and the task-specific effect size for each preprocessing configuration was only formally assessed for the wild-type GA-PD data of each task. For the wild-type POP configuration data of the LTP task no significant correlation was found between the strength of the effect size and either of the data quality measures. This would suggest that the data quality measures cannot be used to predict the preprocessing strategy that will generate the best statistical results. However, as the validity of the GA-PD signal analysis results for the LTP are likely to have been biased by a poor signal estimation approach, as previously discussed, this may explain the lack of correlation. When visual assessment of the significant LTP task effect sizes for the different preprocessing configuration datasets was performed for the results of the ICA-based techniques there appeared to be a positive correlation between the effect size and the SNR. For instance, configuration 7 displayed the largest SNR for both exclusion criteria datasets and went on to produce the largest significant N1b potentiation effect size of any of the datasets analysed by either of the ICA-based signal analysis approaches. Furthermore, significant correlations between the IDAEP task-specific effect size and the SV and SNR measures were observed for the GA-PD analysed, wild-type POP datasets. Although neither measure was able to accurately predict which preprocessing configuration would generate the largest effect size they do show some promise for selecting potential POP configuration subsets for further analysis.

This therefore suggests that the data quality assessment measures performed following data preprocessing may actually be useful in determining optimal datasets for subsequent signal analysis, which consequently could enable future implementations of the preprocessing and signal analysis pipelines to forego the repetitive signal analysis procedures that were required for the exploratory analysis of all of the different POP configurations.

The ability to determine the optimal preprocessing configuration for further analysis based upon post-preprocessing data quality measures such as SNR, may be achieved in future work by investigating the relationship and interactions between multiple preprocessing outcome measures and the associated signal analysis results. For instance, the interaction between the number of participants within the final preprocessed dataset (i.e. post participant exclusion) and statistics related to the SNRs and BVs of each participant within the dataset may all help to better predict the most

appropriate dataset for signal analysis through better representing the key factors that are likely to influence the task-specific statistical tests. Furthermore, it is worth noting that the SNR and SV indexes within this thesis were both calculated using the entire post-stimulus timeframe (i.e. 0ms to 500ms) and as such its relation to the data quality and signal strength of each task-specific SOI is somewhat overgeneralised and therefore could be improved by refining the calculation to a timeframe that more directly relates to the task-specific signal under investigation.

5.2. Limitations

Some of the results obtained from the many different POP and SAP pipeline configurations and analyses were difficult to interpret with regards to understanding how preprocessing parameters and signal analysis techniques may interact and influence task outcomes. An example of this can observed within the LTP task results where a specific POP configuration dataset was found to be significant when IAPD signal analysis was used but not when either of the ICA-based techniques were performed on the same dataset. A further example can also be seen within the IDAEP task data were signal analysis by COIPD would generally result in poor task outcomes however under very specific data preprocessing conditions this signal analysis technique was observed to produce the largest task-specific effect size for the task. Unique instances such as this are difficult to interpret as it is hard to say whether these results represent a true interaction between the type of signal analysis and the unique combination of preprocessing parameters or if it is actually just a by-product of preprocessing induced changes in the participant sample. For instance, the IDAEP dataset produced by ExCrit.BV_{max} POP configuration 15 was unique in that one particular participant was excluded from this dataset that was present within all of the other preprocessing dataset variants. Therefore, if this participant's data was always decomposed poorly by the ICA-based techniques then its exclusion in this instance may be what actually caused the improvement in the COIPD results rather than the specific selection of preprocessing parameters. However, in the case of IDAEP analysis the uncertainty surrounding accurate P2 peak detection will have introduced an element of human error and subjectivity into the results. Therefore, two different approaches could have been used to help validate the peak detection classification measurements. The first would have been to have a second researcher perform peak detection in parallel on the same datasets for later comparison. Alternatively, if the task had of consisted of a greater

number of trials per condition then a split-half analysis technique could have been used to assess the stability of the peak measurements and the results.

Assessing the success of the preprocessing parameters and the signal analysis techniques based upon improvements in the effect size of task-specific outcomes has some inherent limitations that should be discussed. A prime example of how basing the effectiveness of the analysis strategy solely upon the size of the task-specific effect size can be misleading can be seen in the LTP task results that were analysed by the GA-PD approach. In these results the generalised mean amplitude windowing methodology was believed to be causing the effect size of the task to be artificially inflated. As this was likely caused by the N1b amplitude measurement being biased by the activity of the N1 primary visual evoked response, as a result of the imprecision of the N1b mean amplitude window, this highlights how larger effect sizes might not always directly represent better signal estimation.

When talking about the effect sizes of the results from the different analysis configurations for each task the ASF r^2 and the N1b potentiation Cohen d relate to very different statistical metrics. Analysis of the LTP task involves assessing the change in amplitude of the N1b component before and after a tetanising stimulus condition. The reported effect sizes for this task therefore represent both the magnitude of the change as well as the amount of variance within the change. However, analysis of the IDAEP task focused upon assessing the fit of the ASF (i.e. the relationship between the intensity of an auditory stimulus and the amplitude difference between the N1 and P2 auditory evoked potentials) to either a linear or polynomial model. Therefore, if the true ASF of the neuronal sources generating the auditory evoked responses does not fit a perfectly linear or polynomial relationship then using the effect size to determine the accuracy of the signal estimation and isolation could potentially be flawed. However, if we assume that the ASF is linear, then the use of the ASF effect size for the assessment of signal estimation/isolation accuracy could have been improved upon by altering the parameters of the linear fixed effects model.

Within this thesis, the linear model has been established based upon defining the slope of the ASF as a linear fixed effects model with the inclusion of a participant grouping variable to define an ASF with a random intercept for each participant but a shared common gradient. Subsequent likelihood ratio tests, between this linear model and a simple constant (zero-gradient) model, therefore essentially assessed how well each

participant's data fitted to a common ASF gradient. A better assessment of signal isolation would have been achieved if the linear model was designed to allow each participant to have a unique ASF slope and intercept. The likelihood ratio test between the linear and constant model would then directly assess how well each individual participants ASF fitted to a linear relationship with a personalised gradient. If the underlying neuronal signals responsible for the ASF do truly follow a linear relationship, then any improvements in artefact rejection and signal isolation that could be achieved by any of the various analysis strategies would then have directly improved this measure of participant ASF linearity. The only caveat being that this approach would not differentiate between participants that display the expected positive ASF gradient and any participants that display a negative ASF gradient, with each situation providing an equal contribution to the likelihood ratio test of linearity.

5.3. Future directions

The design process of the POP and SAP pipelines presented within this thesis primarily focused upon establishing a predominately automated analysis strategy that would allow for the assessment of both simple and complex signal analysis techniques. The selection of the specific preprocessing mechanisms and the tailored design of the signal analysis methodologies were all chosen to achieve a gradual increase in analytical complexity through the different POP and SAP configurations. Therefore, there were alternative and more advanced artefact removal and signal isolation techniques that could have been included within the analysis pipelines, that would likely have improved the task-specific outcomes, however these were not taken advantage of to limit the overall complexity of the analysis process. For instance OAR could have been improved through the use of hybrid BSS approaches such as REGICA (Klados *et al.*, 2009) or wICA (Castellanos and Makarov, 2006b). Furthermore, the signal decomposition of the ICA-based methodologies could have been improved by first detecting and removing any noisy or outlier EEG channels before computing the ICA followed by reintroducing the excluded channels back into the data by interpolation. In this regard the PREP pipeline designed by Bigdely-Shamlo et al has many robust detection and interpolation algorithms that could be incorporated into future versions of the POP. Using these algorithms the same team also demonstrated how identifying and excluding bad channels is necessary for maintaining signal quality when re-referencing to an average reference (Bigdely-Shamlo *et al.*, 2015). Although using an averaged reference will not

change the effectiveness of the ICA decomposition, its use can however enable the fitting of source dipoles to the ICA components which can go on to provide an additional useful metric for study interpretation and component clustering (Onton *et al.*, 2006; Milne *et al.*, 2009; Ponomarev *et al.*, 2010b; Makeig and Onton, 2011; Delorme *et al.*, 2012).

The component selection and assessment process utilised within this thesis was simply based upon identifying the component with the highest PVAF measure at a predetermined electrode location of interest. However, using more complicated component clustering techniques, such as a multiple measure approach that can assess any combination of various event-related dynamics (e.g. dipole location, spectral power, ERSP, inter-trial coherence, component cross-coherence, ERP contribution and scalp topography) (Onton *et al.*, 2006; Gramann *et al.*, 2010), could aid in objectively identifying similar components across large high-density EEG datasets.

In addition to testing the effectiveness of previously designed analysis strategies, the introduction of the BV_{max} participant exclusion criteria approach was also tested within this thesis and appears to offer a viable method for objectively excluding participants. As the technique is based upon directly assessing participant conditional average data quality it was expected to inherently guarantee a better quality of dataset than that which could be achieved through simply assuming data quality will be sufficient if a minimum number of epochs were present within each conditional average. Further work should therefore be conducted to investigate how this approach might be improved as well as whether a general maximum BV threshold could be used instead of individually tailoring the threshold based upon calculating an estimate of the size of the task-specific changes in the SOI.

The different ERP analysis strategies that have been presented within this thesis only represent a small selection of the types of techniques that have been developed and used within the EEG community over the past 50 years. Although the comparatively newer ICA-based techniques have only been shown to perform marginally better than the trial-averaging signal analysis approaches, more modern advances in ERP analysis promise to push the boundaries of what has been previously possible within ERP research. For instance Burns *et al.* (2013) have recently shown how a regression-based signal analysis technique can be used for segregating overlapping evoked response data in studies were stimuli are presented in close succession. Subsequent comparative analysis of this approach against the standard trial-averaging approach also showed that the

regression-based ERP technique accounted for more event-related signal variance and that subsequent ICA analysis of the regressed data improved the signal decomposition and isolation. Furthermore, Vossen *et al.* (2011) have shown that using mixed regression-based statistical analysis instead of the standard approaches, such as univariate or repeated measures ANOVA, can improve the assessment of single-trial data as well as offer more ways to investigate within and between subject effects.

Single trial analysis is one area of ERP research that is currently gaining a lot of interest within the EEG research community. Within this approach, BSS techniques are used to find singular or groups of independent component/s that relate to an isolated SOI. These isolated source activations are then back-projected to the level of the scalp to allow the assessment of task-specific effects in relation to single trial ERP dynamics such as trial order and response time (Jung *et al.*, 2001; Delorme *et al.*, 2002; Makeig *et al.*, 2004a; Makeig and Onton, 2011; Stewart *et al.*, 2014). Furthermore, Delorme et al. have recently demonstrated a new concept for visualising and statistically assessing the variability in event-related single trials using a grand ERP-image averaging methodology that extends upon the standard ERP-image plotting methodology that has previously been used to view the data from a single recording (Delorme *et al.*, 2015). Therefore, there are many ways in which the POP and SAP pipelines could be adapted and improved upon in the future to help test the effectiveness of these and other techniques.

In this regard, it would be interesting to see how these different preprocessing and signal analysis strategies could be used to improve the clinical utility of EEG-based outcome measures. For instance, as various clinical populations are known to be more likely to produce EEG data with greater proportions of non-neuronal artefacts (Gasser *et al.*, 1992b; Sponheim *et al.*, 1994; Debener *et al.*, 2000; Jung *et al.*, 2000c), these studies would likely greatly benefit from a preprocessing strategy that incorporated a degree of optimisation. In this regard, Cassini et. al. (2014) have already demonstrated the positive affect that various automated artefact removal techniques can have upon the accuracy of an EEG-based Alzeimer's disease diagnostic tool. However, in addition to offering a methodology for improving artefact rejection, using an ERP analysis optimisation pipeline approach such as that which has been developed within this thesis, can also benefit the statistical impact of EEG-related clinical research through improvements in signal isolation and characterisation. Careful consideration and development of the entire analysis pipeline in this way should therefore improve the

detection accuracy of any potential biomarkers of diagnosis, prognosis and/or treatment response that an ERP-based technique may uncover.

5.4. Summary

ICA-based signal analysis techniques such as COIPD can improve task outcomes, however to fully leverage the primary benefits of BSS within these techniques careful and considered task design is required. On the other hand, when clearly defined and stable ERP components are going to be studied then the additional computational complexity and increased analytical workload required by ICA-based techniques may be avoided in favour of simpler trial-averaging ERP approaches such as GAPD and IAPD. However, if the source of the task-specific SOI to be analysed is in close proximity to other large event-related neuronal sources with similar latencies, then GAPD should be avoided in favour of IAPD or if possible COIPD with a high-density EEG recording montage and a specific task design that maximises any unique source distribution properties of the SOI to help isolate it from the neighbouring neuronal activity sources.

For all of the different signal analysis approaches, data preprocessing can clearly be seen to be universally important for improving the data outcomes from these approaches. However, the relative sensitivity of each approach to the various soft POP parameters is still largely unknown. Furthermore, of the multiple preprocessing strategies that were tested for each task no one configuration of preprocessing parameters or specific participant exclusion approach was found to consistently produce the best results across the different signal analysis techniques or tasks. This therefore demonstrates the benefits of exploring and tailoring different preprocessing strategies for each and every task and dataset.

Furthermore, given that the novel BV_{max} participant exclusion criteria approach was shown to be able to offer a viable method for objectively excluding participants, this is one such preprocessing approach that should be explored further in future preprocessing analysis strategies.

Chapter 6. Appendices

6.1. Supplementary Analysis: ICA algorithm training approaches

Training the ICA algorithm upon the concatenated trial-by-trial event-related data is strongly recommended by Vigario and Oja (2008) due the fact that it grants the analysis access to a wealth of data relating to inter-trial differences that would otherwise be lost in the averaging process. Using the trial-by-trial data therefore allows the algorithm to potentially find temporally independent overlapping components that relate to trial-by-trial differences. The benefit of using concatenated conditional average data for ICA training, however, is that it may help to focus the ICA training upon the event specific activity, i.e. minimizing the involvement of non-time- or phase-locked neural or artefactual processes (Makeig 1999). However, although the signal quality of conditional average data can therefore be viewed as 'cleaner', the major drawback to using conditional average data is the drastically reduced amount of datapoints available for ICA training.

As previously mentioned, when Bell and Sejnowski (1995) designed the INFOMAX ICA algorithm it was recommended that it should be used with a minimum data size of 3 times the number of channels squared for a full rank analysis (i.e. decomposing the EEG data into as many components as there are channels in the data). For the studies in this thesis this would give a minimum recommended data size of 2883 datapoints ($3*31^2$). The datapoints available for ICA training under both the conditional average approach and the trial-by-trial approach for both the IDAEP and LTP tasks were calculated as follows:

Epoch (-200ms to 500ms @ 500 Hz) = 350 datapoints

Conditional-Average Data

IDAEP (5 conditions) = 350 * 5 = 1750 datapoints

LTP (4 conditions) = 350 * 4 = 1400 datapoints

Trial-By-Trial Data

IDAEP (At least 20 epochs per condition) = $350 * 5 * 20_{min}$ = 35000 datapoints

LTP (At least 20 epochs per condition) = $350 * 4 * 20_{min}$ = 28000 datapoints

Comparing the minimum number of datapoints, as calculated and recommended by Bell and Sejnowski, with the number of available datapoints for the two tasks and two approaches it was clear that under the trial-by-trial approach the ICA algorithm would have a sufficient amount of data for training. Furthermore, even when following the more recent and stricter recommendation of 20 times the squared number of channels (Onton *et al.*, 2006) the trial-by-trial data available for both tasks still exceeded the minimum number of datapoints recommended for training. However, when assessing the amount of data that would be available under the conditional-average approach both tasks were found to fail to meet even the more lenient of the two minimum datapoint recommendations.

As the minimum data recommendations were not met for the conditional-average datasets of either task it could have been argued that for this reason alone the trial-by-trial approach should have been selected for further analysis. However, although it has been shown that increasing the amount of data that is available for ICA training improves the general reliability of the signal decomposition by increasing the number of reliable ICA components (Onton *et al.*, 2006; Groppe *et al.*, 2009), it has also has also been noted that the most physiologically plausible components were always the most reliable (Groppe *et al.*, 2009). Therefore, given the fact that the ICA-based signal analysis approaches presented within this thesis were designed to exclusively focus upon just a single event-related component, it was hypothesized that the conditional-average approach may still be able to produce a reliable component-of-interest despite the supposed lack of sufficient datapoints for typical ICA training.

ICA training was performed using both approaches and the task-specific significance levels and effect size statistics were calculated for a subset of the pipeline configurations for both tasks as in the main body of analysis within this thesis. During peak detection, it was very evident that the ICA components from the conditional-average datasets typically contained a greater proportion of noise as compared to the ICA components of the trial-by-trial approach (Figure 26 and Figure 27). The additional noise within the conditional-average approach was further evidenced in the significant reduction in PVAF levels observed for this approaches COIs (Table 24). Despite the apparent impairment in signal isolation, the conditional-average approach was still able to demonstrate significant task effects for some of the POP configurations for each of the ICA signal analysis approaches. However, for the majority of datasets the trial-by-trial

approach produced the greater effect-sizes of the two ICA training approaches (Table 25 and Table 26). The only instances where the conditional-average approach produced greater significant effect sizes was when signal analysis was performed via ICA-PD. The fact that this is only observed during ICA-PD analysis and not for COIPD analysis of the same POP configuration data implies that these results simply represent a biased estimation of task-specific SOI due to poor peak latency classification rather than improved signal isolation. Therefore, based upon the general reliability and size of the significant effects sizes produced by the trial-by-trial ICA training approach within this supplementary analysis the results from the trial-by-trial ICA training methodology were selected for inclusion and interpretation within the main body of work.

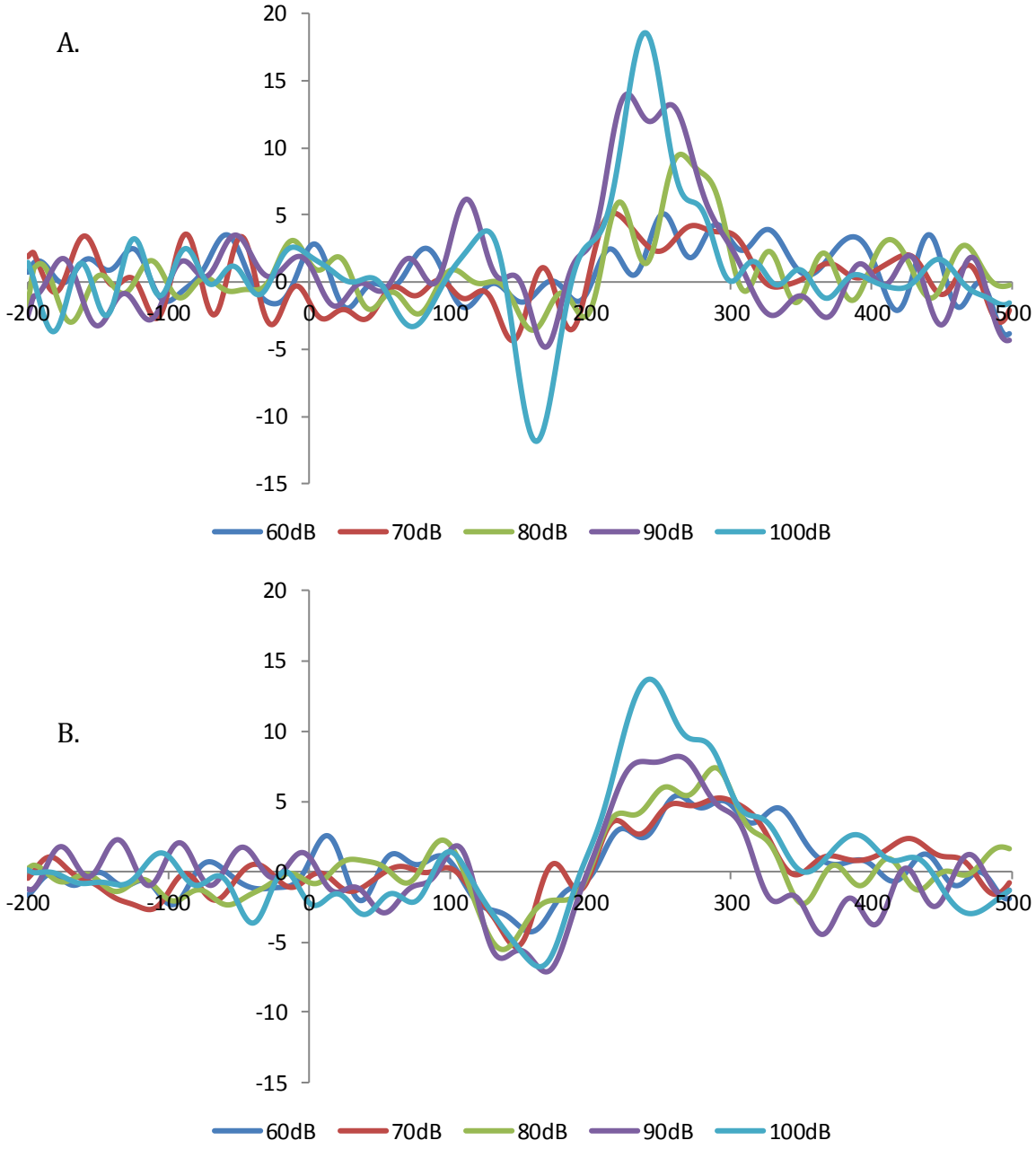

Figure 26. IDAEP task: Back-projected component of interest (COI) activity for a single participant for the conditional-average and trial-by-trial ICA training approaches (A and B respectively). The data displayed was from participant number 27 and was preprocessed under POP configuration 13 and has been selected as a general example of the ICA component data quality at the individual participant level. Note that the ICA component identified by the trial-by-trial approach accounts for a greater amount of the event-related signal variance than the conditional average approach (PVAF (conditional-average) = 68.76, PVAF (trial-by-trial) = 86.59)

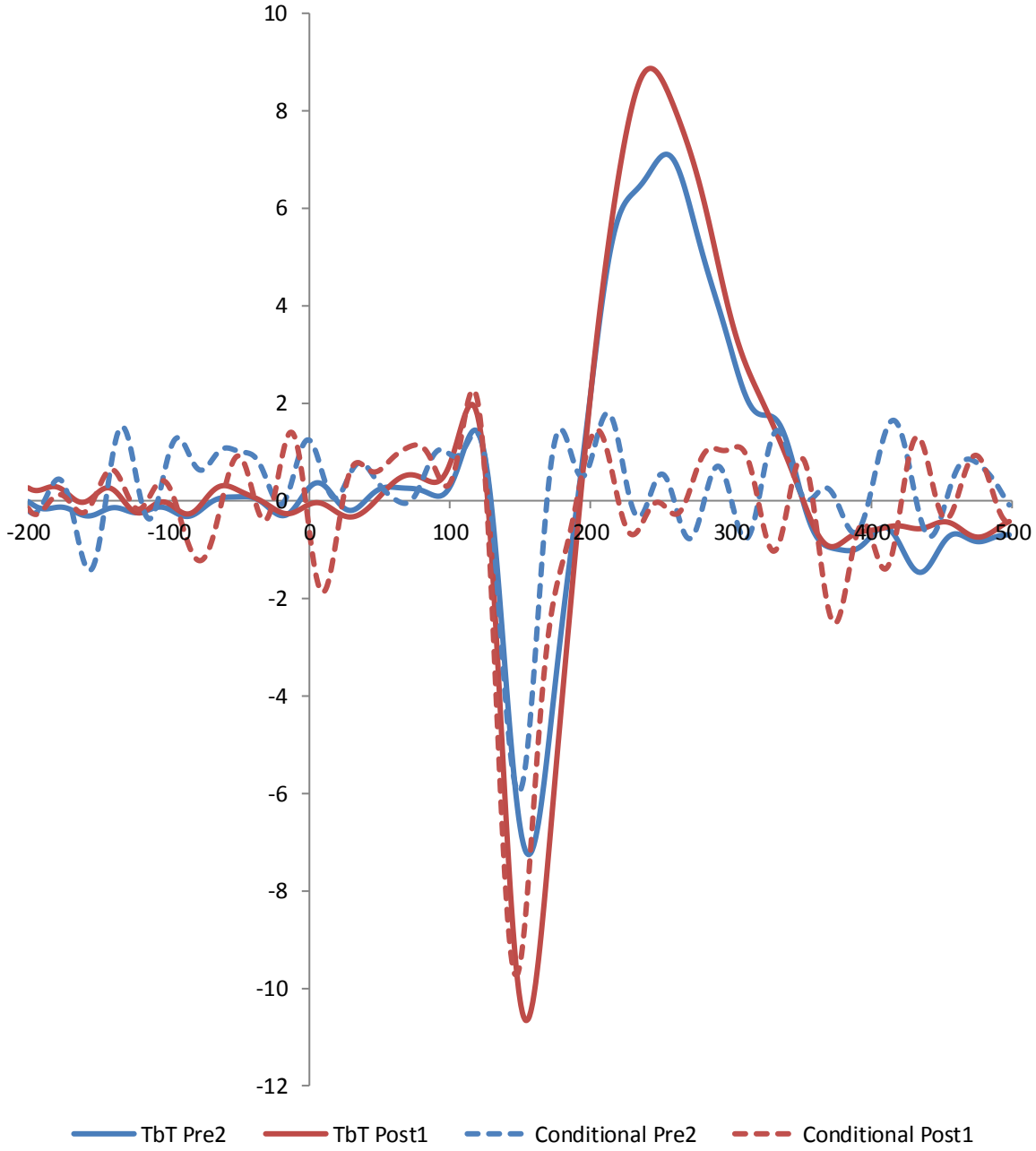

Figure 27. LTP task: Back-projected component of interest (COI) activity for a single participant for the conditional-average and trial-by-trial ICA training approaches (dashed and solid lines respectively). The data displayed was from participant number 30 and was preprocessed under POP configuration 7 and has been selected as a general example of the ICA component data quality at the individual participant level. Note that the ICA component identified by the trial-by-trial approach accounts for a greater amount of the event-related signal variance than the conditional average approach (PVAF (conditional-average) = 46.29, PVAF (trial-by-trial) = 95.76)

Table 24. Percentage variance accounted for (PVAF) descriptive statistics and statistical comparisons for the COIs of the trial-by-trial and conditional-average ICA training approaches, for each task and a subset of preprocessing strategies (A-D). Note that the PVAFs of trial-by-trial COIs were all significantly larger than the PVAFs of the conditional-average approach data.

A. IDAEP: PVAFs ($ExCrit.Epoch_{min}$)							
POP Configuration			Trial-by-Trial		Conditional Average		T-test
Number	Type	N	PVAF (mean)	(s.d)	PVAF (mean)	(s.d)	p-value
11	Gamma	24	58.76	21.56	42.01	13.31	2.64E-04
13	Gamma	24	60.89	19.89	44.75	17.27	2.54E-03
15	Gamma	20	58.05	15.55	38.81	14.01	6.29E-04

B. IDAEP: PVAFs ($ExCrit.BV_{max}$)							
POP Configuration			Trial-by-Trial		Conditional Average		T-test
Number	Type	N	PVAF (mean)	(s.d)	PVAF (mean)	(s.d)	p-value
11	Gamma	21	60.24	20.71	41.79	12.77	2.68E-04
13	Gamma	21	65.12	18.07	45.66	18.17	1.32E-03
15	Gamma	19	58.97	15.40	39.77	12.91	3.00E-04

C. LTP: PVAFs ($ExCrit.Epoch_{min}$)							
POP Configuration			Trial-by-Trial		Conditional Average		T-test
Number	Type	N	PVAF (mean)	(s.d)	PVAF (mean)	(s.d)	p-value
7	Beta	24	81.34	15.11	50.80	13.44	1.21E-07
11	Gamma	28	81.07	14.26	49.33	15.59	3.20E-08
13	Gamma	28	79.99	15.93	53.70	11.89	4.13E-08
15	Gamma	27	74.38	21.44	51.47	12.37	2.97E-04
16	Gamma	26	78.02	19.89	47.41	15.06	1.16E-08

D. LTP: PVAFs ($ExCrit.BV_{max}$)							
POP Configuration			Trial-by-Trial		Conditional Average		T-test
Number	Type	N	PVAF (mean)	(s.d)	PVAF (mean)	(s.d)	p-value
7	Beta	18	81.13	13.55	52.57	14.89	9.75E-06
11	Gamma	22	82.07	13.84	49.56	15.05	5.67E-07
13	Gamma	22	82.55	13.79	54.46	10.43	3.02E-08
15	Gamma	21	74.74	20.88	51.66	11.97	6.26E-04
16	Gamma	21	75.36	20.91	50.20	15.23	4.47E-07

Table 25. IDAEP task-specific outcome statistics and effects sizes for the trial-by-trial and conditional-average data for a subset of preprocessing strategies and the two ICA-based signal analysis techniques (A-D). Where the largest significant task-specific outcomes for either of the ICA algorithm training approaches is highlighted in red.

A. IDAEP: ICA-PD ($ExCrit.Epoch_{min}$)						
POP Configuration			Trial-by-trial		Conditional Average	
Number	Type	N	L vs C p-value	r^2	L vs C p-value	r^2
11	Gamma	24	0.00E+00	0.73	3.11E-15	0.62
13	Gamma	24	0.00E+00	0.70	0.00E+00	0.74
15	Gamma	20	0.00E+00	0.74	8.19E-10	0.71

B. IDAEP: ICA-PD ($ExCrit.BV_{max}$)						
POP Configuration			Trial-by-Trial		Conditional Average	
Number	Type	N	L vs C p-value	r^2	L vs C p-value	r^2
11	Gamma	21	0.00E+00	0.72	5.32E-13	0.59
13	Gamma	21	0.00E+00	0.70	0.00E+00	0.77
15	Gamma	19	0.00E+00	0.75	1.53E-10	0.74

C. IDAEP: COIPD ($ExCrit.Epoch_{min}$)						
POP Configuration			Trial-by-trial		Conditional Average	
Number	Type	N	L vs C p-value	r^2	L vs C p-value	r^2
11	Gamma	24	1.13E-12	0.63	2.40E-12	0.51
13	Gamma	24	1.68E-14	0.67	0.00E+00	0.62
15	Gamma	20	1.95E-11	0.73	5.76E-11	0.45

D. IDAEP: COIPD ($ExCrit.BV_{max}$)						
POP Configuration			Trial-by-Trial		Conditional Average	
Number	Type	N	L vs C p-value	r^2	L vs C p-value	r^2
11	Gamma	21	9.44E-12	0.60	7.34E-12	0.53
13	Gamma	21	1.22E-15	0.69	3.00E-15	0.61
15	Gamma	19	3.89E-15	0.82	7.65E-14	0.56

Table 26. LTP task-specific outcome statistics and effects sizes for the trial-by-trial and conditional-average data for a subset of preprocessing strategies and the two ICA-based signal analysis techniques (A-D). Where the largest significant task-specific outcomes for either of the ICA algorithm training approaches is highlighted in red.

A. LTP: ICA-PD ($ExCrit.Epoch_{min}$)						
POP Configuration			Trial-by-Trial		Conditional Average	
Number	Type	N	p-value	Cohen's d	p-value	Cohen's d
7	Beta	24	2.55E-02	-0.18	3.06E-01	-0.15
11	Gamma	28	9.95E-02	-0.12	2.54E-01	-0.10
13	Gamma	28	3.58E-02	-0.16	1.88E-01	-0.21
15	Gamma	27	5.25E-02	-0.13	1.48E-02	-0.28
16	Gamma	26	1.24E-01	-0.12	2.16E-02	-0.31

B. LTP: ICA-PD ($ExCrit.BV_{max}$)						
POP Configuration			Trial-by-Trial		Conditional Average	
Number	Type	N	p-value	Cohen's d	p-value	Cohen's d
7	Beta	18	8.31E-03	-0.23	4.16E-01	-0.16
11	Gamma	22	2.19E-02	-0.20	8.29E-02	-0.17
13	Gamma	22	4.88E-02	-0.20	5.70E-01	-0.09
15	Gamma	21	1.28E-01	-0.14	1.23E-02	-0.41
16	Gamma	21	6.98E-02	-0.18	6.66E-02	-0.33

C.LTP: COIPD ($ExCrit.Epoch_{min}$)						
POP Configuration			Trial-by-Trial		Conditional Average	
Number	Type	N	p-value	Cohen's d	p-value	Cohen's d
7	Beta	24	3.20E-02	-0.18	1.37E-01	-0.53
11	Gamma	28	4.07E-02	-0.14	8.51E-01	-0.08
13	Gamma	28	1.29E-02	-0.17	4.72E-01	-0.11
15	Gamma	27	1.01E-01	-0.09	1.27E-01	-0.14
16	Gamma	26	1.92E-01	-0.08	5.27E-01	-0.09

D. LTP: COIPD ($ExCrit.BV_{max}$)						
POP Configuration			Trial-by-Trial		Conditional Average	
Number	Type	N	p-value	Cohen's d	p-value	Cohen's d
7	Beta	18	1.91E-02	-0.25	6.92E-01	-0.08
11	Gamma	22	1.70E-02	-0.23	2.56E-02	-0.20
13	Gamma	22	9.15E-03	-0.24	2.36E-01	-0.17
15	Gamma	21	2.63E-01	-0.07	1.42E-01	-0.13
16	Gamma	21	1.60E-01	-0.12	1.42E-01	-0.22

6.2. Supplementary Data: Median ASF statistics for the IDAEP task

Table 27. GA-PD and IAPD ExCrit.Epoch$_{min}$ median ASF statistics (IDAEP task)

POP Configuration			GA-PD		IAPD	
Number	Type	N	p-value	Cohen d	p-value	Cohen d
1	Raw	1	n/a	n/a	n/a	n/a
2	Alpha	1	n/a	n/a	n/a	n/a
3	Beta	1	n/a	n/a	n/a	n/a
4	Beta	1	n/a	n/a	n/a	n/a
5	Beta	7	2.63E-03	0.6401	2.41E-03	1.8959
6	Beta	5	1.87E-03	0.6818	3.02E-02	1.4708
7	Beta	9	2.60E-04	0.8128	5.82E-04	1.8295
8	Beta	5	1.31E-03	0.7099	4.67E-02	1.2712
9	Beta	10	7.93E-04	0.7306	2.08E-04	1.8902
10	Beta	5	1.97E-03	0.6779	5.71E-02	1.1839
11	Gamma	24	2.22E-07	1.3377	4.72E-08	1.6238
12	Gamma	25	1.00E-07	1.4001	4.66E-08	1.5650
13	Gamma	24	1.90E-08	1.5335	7.24E-08	1.5835
14	Gamma	25	3.16E-08	1.4923	2.24E-08	1.6313
15	Gamma	20	1.44E-05	1.0242	5.00E-07	1.6600
16	Gamma	21	7.87E-07	1.2406	3.94E-07	1.6109

Table 28. ICA-PD and COIPD (trial-by-trial approach) ExCrit.Epoch$_{min}$ median ASF statistics (IDAEP task). Where the ICA algorithm was trained upon concatenated trial-by-trial data.

POP Configuration			ICA-PD (Trial-by-trial)		COIPD (Trial-by-trial)	
Number	Type	N	p-value	Cohen d	p-value	Cohen d
11	Gamma	24	1.61E-07	1.5095	1.56E-04	0.9216
13	Gamma	24	8.30E-08	1.5708	8.93E-05	0.9673
15	Gamma	20	1.11E-06	1.5688	2.54E-04	1.0027

Table 29. ICA-PD and COIPD (conditional average approach) ExCrit.Epoch$_{min}$ median ASF statistics (IDAEP task). Where the ICA algorithm was trained upon concatenated conditional average data.

POP Configuration			ICA-PD (Conditional avg.)		COIPD (Conditional acg.)	
Number	Type	N	p-value	Cohen d	p-value	Cohen d
11	Gamma	24	6.64E-06	1.1830	8.28E-06	1.1644
13	Gamma	24	1.64E-06	1.3026	8.93E-08	1.5640
15	Gamma	20	3.43E-04	0.9732	1.20E-05	1.3112

Table 30. GA-PD and IAPD ExCrit.BV$_{max}$ median ASF statistics (IDAEP task)

POP Configuration			GA-PD		IAPD	
Number	Type	N	p-value	Cohen d	p-value	Cohen d
1	Raw	4	5.70E-02	0.3913	1.20E-02	2.7357
2	Alpha	4	4.45E-02	0.4150	3.29E-02	1.8803
3	Beta	3	3.07E-02	0.4492	6.44E-02	2.1635
4	Beta	4	1.78E-02	0.4975	2.47E-02	2.0975
5	Beta	12	1.24E-04	0.8674	3.65E-04	1.4614
6	Beta	7	4.75E-04	0.7877	2.46E-02	1.1270
7	Beta	15	1.69E-04	0.8444	1.72E-04	1.3083
8	Beta	7	2.21E-03	0.6686	1.03E-02	1.3917
9	Beta	15	1.16E-04	0.8717	5.14E-04	1.1580
10	Beta	8	1.91E-04	0.8568	1.38E-02	1.1535
11	Gamma	21	1.02E-06	1.2207	8.71E-07	1.5261
12	Gamma	21	7.13E-07	1.2481	1.30E-06	1.4844
13	Gamma	21	2.53E-07	1.3276	4.78E-07	1.5901
14	Gamma	21	8.30E-07	1.2365	5.44E-07	1.5763
15	Gamma	19	2.66E-05	0.9793	1.04E-06	1.6543
16	Gamma	23	2.27E-07	1.3361	2.71E-07	1.5184

Table 31. ICA-PD and COIPD (trial-by-trial approach) ExCrit.BV$_{max}$ median ASF statistics (IDAEP task). Where the ICA algorithm was trained upon concatenated trial-by-trial data.

POP Configuration			ICA-PD (Trial-by-trial)		COIPD (Trial-by-trial)	
Number	Type	N	p-value	Cohen d	p-value	Cohen d
11	Gamma	21	1.67E-06	1.4581	3.27E-04	0.9443
13	Gamma	21	5.50E-07	1.5750	4.70E-05	1.1272
15	Gamma	19	4.59E-07	1.7554	3.14E-06	1.5225

Table 32. ICA-PD and COIPD (conditional average approach) ExCrit.BV$_{max}$ median ASF statistics (IDAEP task). Where the ICA algorithm was trained upon concatenated trial-by-trial data.

POP Configuration			ICA-PD (Conditional avg.)		COIPD (Conditional avg.)	
Number	Type	N	p-value	Cohen d	p-value	Cohen d
11	Gamma	21	4.86E-05	1.124100	5.69E-05	1.109035
13	Gamma	21	1.12E-06	1.499314	1.12E-06	1.499353
15	Gamma	19	1.79E-04	1.077920	4.05E-06	1.493046

Chapter 7. References

Acunzo, D.J., Mackenzie, G. and van Rossum, M.C. (2012) 'Systematic biases in early ERP and ERF components as a result of high-pass filtering', *J Neurosci Methods*, 209(1), pp. 212-8.

Adler, G. and Adler, J. (1989) 'Influence of stimulus intensity on AEP components in the 80- to 200-millisecond latency range', *Audiology*, 28(6), pp. 316-24.

Ai, G., Sato, N., Singh, B. and Wagatsuma, H. 'Direction and viewing area-sensitive influence of EOG artifacts revealed in the EEG topographic pattern analysis', *Cognitive Neurodynamics*, pp. 1-14.

Alain, C., Woods, D.L. and Covarrubias, D. (1997) 'Activation of duration-sensitive auditory cortical fields in humans', *Electroencephalogr Clin Neurophysiol*, 104(6), pp. 531-9.

Albera, L., Kachenoura, A., Comon, P., Karfoul, A., Wendling, F., Senhadji, L. and Merlet, I. (2012) 'ICA-based EEG denoising: a comparative analysis of fifteen methods', *Bulletin of the Polish Academy of Sciences: Technical Sciences*, 60(3), pp. 407-418.

Amitay, S., Guiraud, J., Sohoglu, E., Zobay, O., Edmonds, B.A., Zhang, Y.X. and Moore, D.R. (2013) 'Human decision making based on variations in internal noise: an EEG study', *PLoS One*, 8(7), p. e68928.

Aurlien, H., Gjerde, I.O., Aarseth, J.H., Eldoen, G., Karlsen, B., Skeidsvoll, H. and Gilhus, N.E. (2004) 'EEG background activity described by a large computerized database', *Clin Neurophysiol*, 115(3), pp. 665-73.

Beagley, H.A. and Knight, J.J. (1967) 'Changes in auditory evoked response with intensity', *J Laryngol Otol*, 81(8), pp. 861-73.

Bell, A.J. and Sejnowski, T.J. (1995) 'An information-maximization approach to blind separation and blind deconvolution', *Neural computation*, 7(6), pp. 1129-1159.

Belouchrani, A., Abed-Meraim, K., Cardoso, J.F. and Moulines, E. (1997) 'A blind source separation technique using second-order statistics', *IEEE Transactions on signal processing*, 45(2), pp. 434-444.

Berg, P. and Scherg, M. (1991) 'Dipole models of eye movements and blinks', *Electroencephalogr Clin Neurophysiol*, 79(1), pp. 36-44.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M. and Robbins, K.A. (2015) 'The PREP pipeline: standardized preprocessing for large-scale EEG analysis', *Frontiers in Neuroinformatics*, 9, p. 16.

Bishop, D.V., Anderson, M., Reid, C. and Fox, A.M. (2011) 'Auditory development between 7 and 11 years: an event-related potential (ERP) study', *PLoS One*, 6(5), p. e18993.
Bokil, H., Andrews, P., Kulkarni, J.E., Mehta, S. and Mitra, P.P. (2010) 'Chronux: a platform for analyzing neural signals', *Journal of neuroscience methods*, 192(1), pp. 146-151.
Boutros, N. (2008) 'Lack of blinding in gating studies', *Schizophrenia research*, 103(1-3), p. 336; author reply 337.
Bruneau, N., Roux, S., Garreau, B. and Lelord, G. (1985) 'Frontal auditory evoked potentials and augmenting-reducing', *Electroencephalogr Clin Neurophysiol*, 62(5), pp. 364-71.
Burns, M.D., Bigdely-Shamlo, N., Smith, N.J., Kreutz-Delgado, K. and Makeig, S. (2013) *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE.
Callaway, E. (2012) *Event-related brain potentials in man*. Elsevier.
Cassani, R., Falk, T.H., Fraga, F.J., Kanda, P.A. and Anghinah, R. (2014) 'The effects of automated artifact removal algorithms on electroencephalography-based Alzheimer's disease diagnosis', *Front Aging Neurosci*, 6, p. 55.
Castellanos, N.P. and Makarov, V.A. (2006a) 'Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis', *Journal of neuroscience methods*, 158(2), pp. 300-312.
Castellanos, N.P. and Makarov, V.A. (2006b) 'Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis', *J Neurosci Methods*, 158(2), pp. 300-12.
Cavus, I., Reinhart, R.M., Roach, B.J., Gueorguieva, R., Teyler, T.J., Clapp, W.C., Ford, J.M., Krystal, J.H. and Mathalon, D.H. (2012) 'Impaired visual cortical plasticity in schizophrenia', *Biol Psychiatry*, 71(6), pp. 512-20.
Chang, W.P., Gavin, W.J. and Davies, P.L. (2012) 'Bandpass filter settings differentially affect measurement of P50 sensory gating in children and adults', *Clin Neurophysiol*, 123(11), pp. 2264-72.
Chen, W.R., Lee, S., Kato, K., Spencer, D.D., Shepherd, G.M. and Williamson, A. (1996) 'Long-term modifications of synaptic efficacy in the human inferior and middle temporal cortex', *Proceedings of the National Academy of Sciences*, 93(15), pp. 8011-8015.
Clapp, W.C., Hamm, J.P., Kirk, I.J. and Teyler, T.J. (2012) 'Translating long-term potentiation from animals to humans: a novel method for noninvasive assessment of cortical plasticity', *Biol Psychiatry*, 71(6), pp. 496-502.

Clapp, W.C., Zaehle, T., Lutz, K., Marcar, V.L., Kirk, I.J., Hamm, J.P., Teyler, T.J., Corballis, M.C. and Jancke, L. (2005) 'Effects of long-term potentiation in the human visual cortex: a functional magnetic resonance imaging study', *Neuroreport*, 16(18), pp. 1977-80.
Coenen, A.M. (1995) 'Neuronal activities underlying the electroencephalogram and evoked potentials of sleeping and waking: implications for information processing', *Neurosci Biobehav Rev*, 19(3), pp. 447-63.
Cohen, J. (1988) 'Statistical power analysis for the behavioral sciences (second ed.)'. New York: Academic Press.
Cohen, J. (1992) 'A power primer', *Psychological bulletin*, 112(1), p. 155.
Collura, T.F. (1993) 'History and evolution of electroencephalographic instruments and techniques', *J Clin Neurophysiol*, 10(4), pp. 476-504.
Cook, E.W. and Miller, G.A. (1992) 'Digital filtering: Background and tutorial for psychophysiologists', *Psychophysiology*, 29(3), pp. 350-362.
Cooke, S.F. and Bear, M.F. (2012) 'Stimulus-selective response plasticity in the visual cortex: an assay for the assessment of pathophysiology and treatment of cognitive impairment associated with psychiatric disorders', *Biol Psychiatry*, 71(6), pp. 487-95.
Corby, J.C. and Kopell, B.S. (1972) 'Differential contributions of blinks and vertical eye movements as artifacts in EEG recording', *Psychophysiology*, 9(6), pp. 640-4.
Crespo-Garcia, M., Atienza, M. and Cantero, J.L. (2008a) 'Muscle artifact removal from human sleep EEG by using independent component analysis', *Ann Biomed Eng*, 36(3), pp. 467-75.
Crespo-Garcia, M., Atienza, M. and Cantero, J.L. (2008b) 'Muscle artifact removal from human sleep EEG by using independent component analysis', *Annals of biomedical engineering*, 36(3), pp. 467-475.
Croft, R.J. and Barry, R.J. (1998) 'EOG correction: a new perspective', *Electroencephalogr Clin Neurophysiol*, 107(6), pp. 387-94.
Croft, R.J. and Barry, R.J. (2000a) 'EOG correction of blinks with saccade coefficients: a test and revision of the aligned-artefact average solution', *Clin Neurophysiol*, 111(3), pp. 444-51.
Croft, R.J. and Barry, R.J. (2000b) 'EOG correction: which regression should we use?', *Psychophysiology*, 37(1), pp. 123-5.
Croft, R.J. and Barry, R.J. (2000c) 'Removal of ocular artifact from the EEG: a review', *Neurophysiol Clin*, 30(1), pp. 5-19.

Croft, R.J., Chandler, J.S., Barry, R.J., Cooper, N.R. and Clarke, A.R. (2005) 'EOG correction: a comparison of four methods', *Psychophysiology*, 42(1), pp. 16-24.
Crowley, K.E. and Colrain, I.M. (2004) 'A review of the evidence for P2 being an independent component process: age, sleep and modality', *Clin Neurophysiol*, 115(4), pp. 732-44.
Csibra, G., Tucker, L.A. and Johnson, M.H. (2001) 'Differential frontal cortex activation before anticipatory and reactive saccades in infants', *Infancy*, 2(2), pp. 159-174.
Davis, H., Bowers, C. and Hirsh, S.K. (1968) 'Relations of the human vertex potential to acoustic input: Loudness and masking', *The Journal of the Acoustical Society of America*, 43(3), pp. 431-438.
Davis, H. and Zerlin, S. (1966) 'Acoustic relations of the human vertex potential', *The Journal of the Acoustical Society of America*, 39(1), pp. 109-116.
Davis, P.A. (1939) 'Effects of acoustic stimuli on the waking human brain', *Journal of Neurophysiology*, 2(6), pp. 494-499.
De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W. and Van Huffel, S. (2006) 'Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram', *IEEE transactions on Biomedical Engineering*, 53(12), pp. 2583-2587.
de Gobbi Porto, F.H., Fox, A.M., Tusch, E.S., Sorond, F., Mohammed, A.H. and Daffner, K.R. (2015) 'In vivo evidence for neuroplasticity in older adults', *Brain research bulletin*, 114, pp. 56-61.
Debener, S., Beauducel, A., Nessler, D., Brocke, B., Heilemann, H. and Kayser, J. (2000) 'Is resting anterior EEG alpha asymmetry a trait marker for depression? Findings for healthy adults and clinically depressed patients', *Neuropsychobiology*, 41(1), pp. 31-7.
Debener, S., Makeig, S., Delorme, A. and Engel, A.K. (2005) 'What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis', *Cognitive Brain Research*, 22(3), pp. 309-321.
Debnath, L. and Shah, F.A. (2002) *Wavelet transforms and their applications*. Springer.
Delorme, A. and Makeig, S. (2003) 'EEG changes accompanying learned regulation of 12-Hz EEG activity', *IEEE Transactions on neural systems and Rehabilitation engineering*, 11(2), pp. 133-137.

Delorme, A. and Makeig, S. (2004) 'EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis', *Journal of neuroscience methods*, 134(1), pp. 9-21.
Delorme, A., Makeig, S., Fabre-Thorpe, M. and Sejnowski, T. (2002) 'From single-trial EEG to brain area dynamics', *Neurocomputing*, 44, pp. 1057-1064.
Delorme, A., Miyakoshi, M., Jung, T.-P. and Makeig, S. (2015) 'Grand average ERP-image plotting and statistics: A method for comparing variability in event-related single-trial EEG activities across subjects and conditions', *Journal of neuroscience methods*, 250, pp. 3-6.
Delorme, A., Palmer, J., Onton, J., Oostenveld, R. and Makeig, S. (2012) 'Independent EEG sources are dipolar', *PLoS One*, 7(2), p. e30135.
Delorme, A., Sejnowski, T. and Makeig, S. (2007) 'Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis', *Neuroimage*, 34(4), pp. 1443-9.
Di Russo, F., Martínez, A., Sereno, M.I., Pitzalis, S. and Hillyard, S.A. (2002) 'Cortical sources of the early components of the visual evoked potential', *Human brain mapping*, 15(2), pp. 95-111.
Duncan-Johnson, C.C. and Donchin, E. (1979) 'The time constant in P300 recording', *Psychophysiology*, 16(1), pp. 53-55.
Engebretson, A.M. (1963) *A digital computer for analyzing certain bioelectric signals*.
Evans, I.D., Jamieson, G., Croft, R. and Pham, T.T. (2012) 'Empirically validating fully automated EOG artifact correction using independent components analysis', [Online]. Available at: www.frontiersin.org/10.3389/conf.fnhum.2012.208.00034/event_abstract.
Fitzgibbon, S.P., Powers, D.M., Pope, K.J. and Clark, C.R. (2007) 'Removal of EEG noise and artifact using blind source separation', *J Clin Neurophysiol*, 24(3), pp. 232-43.
Gasser, T., Ziegler, P. and Gattaz, W.F. (1992a) 'The deleterious effect of ocular artefacts on the quantitative EEG, and a remedy', *European archives of psychiatry and clinical neuroscience*, 241(6), pp. 352-356.
Gasser, T., Ziegler, P. and Gattaz, W.F. (1992b) 'The deleterious effect of ocular artefacts on the quantitative EEG, and a remedy', *Eur Arch Psychiatry Clin Neurosci*, 241(6), pp. 352-6.
Gómez-Herrero, G. (2007) 'Automatic artifact removal (AAR) toolbox v1. 3 (Release 09.12. 2007) for MATLAB', *Tampere University of Technology*.

Gómez-Herrero, G., De Clercq, W., Anwar, H., Kara, O., Egiazarian, K., Van Huffel, S. and Van Paesschen, W. (2006) *Proceedings of the 7th Nordic Signal Processing Symposium-NORSIG 2006*. IEEE.

Gramann, K., Gwin, J.T., Bigdely-Shamlo, N., Ferris, D.P. and Makeig, S. (2010) 'Visual evoked responses during standing and walking', *Front Hum Neurosci*, 4, p. 202.

Gratton, G., Coles, M.G.H. and Donchin, E. (1983) 'A new method for off-line removal of ocular artifact', *Electroencephalography and clinical neurophysiology*, 55(4), pp. 468-484.

Groppe, D.M., Makeig, S. and Kutas, M. (2009) 'Identifying reliable independent components via split-half comparisons', *Neuroimage*, 45(4), pp. 1199-211.

Hagenmuller, F., Heekeren, K., Meier, M., Theodoridou, A., Walitza, S., Haker, H., Rossler, W. and Kawohl, W. (2016) 'The Loudness Dependence of Auditory Evoked Potentials (LDAEP) in individuals at risk for developing bipolar disorders and schizophrenia', *Clin Neurophysiol*, 127(2), pp. 1342-50.

Hagenmuller, F., Hitz, K., Darvas, F. and Kawohl, W. (2011) 'Determination of the loudness dependence of auditory evoked potentials: single-electrode estimation versus dipole source analysis', *Hum Psychopharmacol*, 26(2), pp. 147-54.

Hajcak, G., Weinberg, A., MacNamara, A. and Foti, D. (2012) 'ERPs and the study of emotion', *The Oxford handbook of event-related potential components*, pp. 441-474.

Harrison, P.J. and Weinberger, D.R. (2005) 'Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence', *Mol Psychiatry*, 10(1), pp. 40-68; image 5.

He, P., Wilson, G. and Russell, C. (2004) 'Removal of ocular artifacts from electro-encephalogram by adaptive filtering', *Medical and biological engineering and computing*, 42(3), pp. 407-412.

He, P., Wilson, G., Russell, C. and Gerschutz, M. (2007) 'Removal of ocular artifacts from the EEG: a comparison between time-domain regression method and adaptive filtering method using simulated data', *Medical & biological engineering & computing*, 45(5), pp. 495-503.

Hegerl, U., Gallinat, J. and Juckel, G. (2001) 'Event-related potentials: do they reflect central serotonergic neurotransmission and do they predict clinical response to serotonin agonists?', *Journal of affective disorders*, 62(1), pp. 93-100.

Hegerl, U. and Juckel, G. (1993) 'Intensity dependence of auditory evoked potentials as an indicator of central serotonergic neurotransmission: a new hypothesis', *Biol Psychiatry*, 33(3), pp. 173-87.

Hegerl, U., Wulff, H. and Muller-Oerlinghausen, B. (1992) 'Intensity dependence of auditory evoked potentials and clinical response to prophylactic lithium medication: a replication study', *Psychiatry Res*, 44(3), pp. 181-90.
Hensch, T., Herold, U., Diers, K., Armbruster, D. and Brocke, B. (2008) 'Reliability of intensity dependence of auditory-evoked potentials', *Clin Neurophysiol*, 119(1), pp. 224-36.
Heynen, A.J. and Bear, M.F. (2001) 'Long-term potentiation of thalamocortical transmission in the adult visual cortex in vivo', *The Journal of Neuroscience*, 21(24), pp. 9801-9813.
Hillyard, S.A. and Galambos, R. (1970) 'Eye movement artifact in the CNV', *Electroencephalogr Clin Neurophysiol*, 28(2), pp. 173-82.
Hoehl, S. and Wahl, S. (2012) 'Recording infant ERP data for cognitive research', *Dev Neuropsychol*, 37(3), pp. 187-209.
Hopf, J.M., Vogel, E., Woodman, G., Heinze, H.J. and Luck, S.J. (2002) 'Localizing visual discrimination processes in time and space', *J Neurophysiol*, 88(4), pp. 2088-95.
Hu, L., Boutros, N.N. and Jansen, B.H. (2009) 'Evoked potential variability', *J Neurosci Methods*, 178(1), pp. 228-36.
Hyvärinen, A., Karhunen, J. and Oja, E. (2004) *Independent component analysis*. John Wiley & Sons.
Ifeachor, E.C. and Jervis, B.W. (2002) *Digital signal processing: a practical approach*. Pearson Education.
Jervis, B.W., Thomlinson, M., Mair, C., Lopez, J.M.L. and Garcia, M.I.B. (1999) 'Residual ocular artefact subsequent to ocular artefact removal from the electroencephalogram', *IEE Proceedings-Science, Measurement and Technology*, 146(6), pp. 293-298.
Jia, X. and Kohn, A. (2011) 'Gamma rhythms in the brain', *PLoS Biol*, 9(4), p. e1001045.
Joyce, C.A., Gorodnitsky, I.F. and Kutas, M. (2004) 'Automatic removal of eye movement and blink artifacts from EEG data using blind component separation', *Psychophysiology*, 41(2), pp. 313-325.
Juckel, G. (2015) 'Serotonin: from sensory processing to schizophrenia using an electrophysiological method', *Behav Brain Res*, 277, pp. 121-4.
Juckel, G., Gallinat, J., Riedel, M., Sokullu, S., Schulz, C., Moller, H.J., Muller, N. and Hegerl, U. (2003) 'Serotonergic dysfunction in schizophrenia assessed by the loudness dependence measure of primary auditory cortex evoked activity', *Schizophr Res*, 64(2-3), pp. 115-24.

Juckel, G., Hegerl, U., Molnár, M., Csépe, V. and Karmos, G. (1999) 'Auditory evoked potentials reflect serotonergic neuronal activity—a study in behaving cats administered drugs acting on 5-HT1A autoreceptors in the dorsal raphe nucleus', *Neuropsychopharmacology*, 21(6), pp. 710-716.

Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M.J., Iragui, V. and Sejnowski, T.J. (1998a) 'Extended ICA removes artifacts from electroencephalographic recordings', *Advances in neural information processing systems*, pp. 894-900.

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M.J., Iragui, V. and Sejnowski, T.J. (2000a) 'Removing electroencephalographic artifacts by blind source separation', *Psychophysiology*, 37(02), pp. 163-178.

Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E. and Sejnowski, T.J. (2000b) 'Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects', *Clinical Neurophysiology*, 111(10), pp. 1745-1758.

Jung, T.P., Humphries, C., Lee, T.W., Makeig, S., McKeown, M.J., Iragui, V. and Sejnowski, T.J. (1998b) *Neural Networks for Signal Processing VIII, 1998. Proceedings of the 1998 IEEE Signal Processing Society Workshop*. IEEE.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E. and Sejnowski, T.J. (2000c) 'Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects', *Clin Neurophysiol*, 111(10), pp. 1745-58.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E. and Sejnowski, T.J. (2001) 'Analysis and visualization of single-trial event-related potentials', *Human brain mapping*, 14(3), pp. 166-185.

Kaatiala, J., Yrttiaho, S., Forssman, L., Perdue, K. and Leppanen, J. (2014) 'A graphical user interface for infant ERP analysis', *Behav Res Methods*, 46(3), pp. 745-57.

Kaskey, G.B., Salzman, L.F., Klorman, R. and Pass, H.L. (1980) 'Relationships between stimulus intensity and amplitude of visual and auditory event related potentials', *Biol Psychol*, 10(2), pp. 115-25.

Kavitha, P.T., Lau, C.T. and Premkumar, A.B. (2007) *Information, Communications & Signal Processing, 2007 6th International Conference on*. IEEE.

Kenemans, J.L. and Kähkönen, S. (2011) 'How human electrophysiology informs psychopharmacology: from bottom-up driven processing to top-down control', *Neuropsychopharmacology*, 36(1), pp. 26-51.

Kierkels, J.J.M., van Boxtel, G.J.M. and Vogten, L.L.M. (2006) 'A model-based objective evaluation of eye movement correction in EEG recordings', *IEEE Transactions on biomedical engineering*, 53(2), pp. 246-253.
Klados, M.A. and Bamidis, P.D. (2016) 'A semi-simulated EEG/EOG dataset for the comparison of EOG artifact rejection techniques', *Data in Brief*, 8, pp. 1004-1006.
Klados, M.A., Bratsas, C., Frantzidis, C., Papadelis, C.L. and Bamidis, P.D. (2010) *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010*. Springer.
Klados, M.A., Papadelis, C., Braun, C. and Bamidis, P.D. (2011) 'REG-ICA: A hybrid methodology combining Blind Source Separation and regression techniques for the rejection of ocular artifacts', *Biomedical Signal Processing and Control*, 6(3), pp. 291-300.
Klados, M.A., Papadelis, C.L. and Bamidis, P.D. (2009) *2009 9th International Conference on Information Technology and Applications in Biomedicine*. IEEE.
Klem, G.H., Luders, H.O., Jasper, H.H. and Elger, C. (1999) 'The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology', *Electroencephalogr Clin Neurophysiol Suppl*, 52, pp. 3-6.
Komatsu, Y., Fujii, K., Maeda, J., Sakaguchi, H. and Toyama, K. (1988) 'Long-term potentiation of synaptic transmission in kitten visual cortex', *Journal of Neurophysiology*, 59(1), pp. 124-141.
Lee, K.S., Park, Y.M. and Lee, S.H. (2012) 'Serotonergic dysfunction in patients with bipolar disorder assessed by the loudness dependence of the auditory evoked potential', *Psychiatry Investig*, 9(3), pp. 298-306.
Lee, T.W., Yu, Y.W., Chen, T.J. and Tsai, S.J. (2005) 'Loudness dependence of the auditory evoked potential and response to antidepressants in Chinese patients with major depression', *J Psychiatry Neurosci*, 30(3), pp. 202-5.
Lin, Y.-P., Wang, Y., Wei, C.-S. and Jung, T.-P. (2013) *International Conference on Human-Computer Interaction*. Springer.
Linka, T., Muller, B.W., Bender, S. and Sartory, G. (2004) 'The intensity dependence of the auditory evoked N1 component as a predictor of response to Citalopram treatment in patients with major depression', *Neurosci Lett*, 367(3), pp. 375-8.
Linka, T., Sartory, G., Bender, S., Gastpar, M. and Muller, B.W. (2007) 'The intensity dependence of auditory ERP components in unmedicated patients with major depression and healthy controls. An analysis of group differences', *J Affect Disord*, 103(1-3), pp. 139-45.

Linka, T., Sartory, G., Gastpar, M., Scherbaum, N. and Muller, B.W. (2009) 'Clinical symptoms of major depression are associated with the intensity dependence of auditory event-related potential components', *Psychiatry Res*, 169(2), pp. 139-43.
Lømo, T. (1966) 'Frequency potentiation of excitatory synaptic activity in the dentate area of the hippocampal formation', *Acta Physiologica*, 68 (Supplement 277)(128).
Luck, S.J. (2005) *An introduction to the event-related potential technique*. MIT press.
Luck, S.J. (2014) *An introduction to the event-related potential technique*. MIT press.
Makeig, S., Bell, A.J., Jung, T.-P. and Sejnowski, T.J. (1996) 'Independent component analysis of electroencephalographic data', *Advances in neural information processing systems*, pp. 145-151.
Makeig, S., Debener, S., Onton, J. and Delorme, A. (2004a) 'Mining event-related brain dynamics', *Trends in cognitive sciences*, 8(5), pp. 204-210.
Makeig, S., Delorme, A., Westerfield, M., Jung, T.-P., Townsend, J., Courchesne, E. and Sejnowski, T.J. (2004b) 'Electroencephalographic brain dynamics following manually responded visual targets', *PLoS Biol*, 2(6), p. e176.
Makeig, S., Jung, T.P., Bell, A.J., Ghahremani, D. and Sejnowski, T.J. (1997) 'Blind separation of auditory event-related brain responses into independent components', *Proc Natl Acad Sci U S A*, 94(20), pp. 10979-84.
Makeig, S. and Onton, J. (2011) 'ERP features and EEG dynamics: an ICA perspective', *Oxford handbook of event-related potential components*.
Makeig, S., Westerfield, M., Jung, T.-P., Covington, J., Townsend, J., Sejnowski, T.J. and Courchesne, E. (1999) 'Functionally independent components of the late positive event-related potential during visual spatial attention', *The journal of neuroscience*, 19(7), pp. 2665-2680.
Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E. and Sejnowski, T.J. (2002) 'Dynamic brain sources of visual evoked responses', *Science*, 295(5555), pp. 690-4.
Mangun, G.R. and Hillyard, S.A. (1990) 'Allocation of visual attention to spatial locations: tradeoff functions for event-related brain potentials and detection performance', *Percept Psychophys*, 47(6), pp. 532-50.
McAllister-Williams, R.H., Anderson, I.M., Finkelmeyer, A., Gallagher, P., Grunze, H.C., Haddad, P.M., Hughes, T., Lloyd, A.J., Mamasoula, C., McColl, E., Pearce, S., Siddiqi, N., Sinha, B.N., Steen, N., Wainwright, J., Winter, F.H., Ferrier, I.N. and Watson, S. (2016) 'Antidepressant augmentation with metyrapone for treatment-resistant depression (the

ADD study): a double-blind, randomised, placebo-controlled trial', *Lancet Psychiatry*, 3(2), pp. 117-27.
McGregor, C.H. (2016) *Plug, socket & voltage by country*. Available at: http://www.worldstandards.eu/electricity/plug-voltage-by-country/.
McNair, N.A., Clapp, W.C., Hamm, J.P., Teyler, T.J., Corballis, M.C. and Kirk, I.J. (2006) 'Spatial frequency-specific potentiation of human visual-evoked potentials', *Neuroreport*, 17(7), pp. 739-41.
Mears, R.P. and Spencer, K.M. (2012) 'Electrophysiological assessment of auditory stimulus-specific plasticity in schizophrenia', *Biol Psychiatry*, 71(6), pp. 503-11.
Mehta, J., Jerger, S., Jerger, J. and Martin, J. (2009) 'Electrophysiological correlates of word comprehension: event-related potential (ERP) and independent component analysis (ICA)', *Int J Audiol*, 48(1), pp. 1-11.
Miller, E.K. and Desimone, R. (1994) 'Parallel neuronal mechanisms for short-term memory', *Science*, 263(5146), pp. 520-2.
Miller, E.K., Li, L. and Desimone, R. (1991) 'A neural mechanism for working and recognition memory in inferior temporal cortex', *Science*, 254(5036), pp. 1377-9.
Milne, E., Scope, A., Pascalis, O., Buckley, D. and Makeig, S. (2009) 'Independent component analysis reveals atypical electroencephalographi c activity during visual perception in individuals with autism', *Biol Psychiatry*, 65(1), pp. 22-30.
Mitra, P.P. and Pesaran, B. (1999) 'Analysis of dynamic brain imaging data', *Biophysical journal*, 76(2), pp. 691-708.
Miyakoshi, M. (2016). Available at: https://sccn.ucsd.edu/wiki/Makoto's_preprocessing_pipeline.
Miyashita, Y. (1993) 'Inferior temporal cortex: where visual perception meets memory', *Annu Rev Neurosci*, 16, pp. 245-63.
Möcks, J., Gasser, T. and Tuan, P.D. (1984) 'Variability of single visual evoked potentials evaluated by two new statistical tests', *Electroencephalography and clinical neurophysiology*, 57(6), pp. 571-580.
Mullen, T. (2012) *CleanLine EEGLAB plugin* [Computer program]. Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC), San Diego, CA (2012).
Naatanen, R. and Picton, T. (1987) 'The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure', *Psychophysiology*, 24(4), pp. 375-425.

Nikolaev, A.R., Jurica, P., Nakatani, C., Plomp, G. and van Leeuwen, C. (2013) 'Visual encoding and fixation target selection in free viewing: presaccadic brain potentials', *Front Syst Neurosci*, 7, p. 26.
Normann, C., Schmitz, D., Furmaier, A., Doing, C. and Bach, M. (2007) 'Long-term plasticity of visually evoked potentials in humans is altered in major depression', *Biol Psychiatry*, 62(5), pp. 373-80.
Ochoa, C.J. and Polich, J. (2000) 'P300 and blink instructions', *Clinical Neurophysiology*, 111(1), pp. 93-98.
Onton, J., Westerfield, M., Townsend, J. and Makeig, S. (2006) 'Imaging human EEG dynamics using independent component analysis', *Neuroscience & Biobehavioral Reviews*, 30(6), pp. 808-822.
Ostermann, J., Uhl, I., Kohler, E., Juckel, G. and Norra, C. (2012) 'The loudness dependence of auditory evoked potentials and effects of psychopathology and psychopharmacotherapy in psychiatric inpatients', *Hum Psychopharmacol*, 27(6), pp. 595-604.
Paiva, T.O., Almeida, P.R., Ferreira-Santos, F., Vieira, J.B., Silveira, C., Chaves, P.L., Barbosa, F. and Marques-Teixeira, J. (2016) 'Similar sound intensity dependence of the N1 and P2 components of the auditory ERP: Averaged and single trial evidence', *Clin Neurophysiol*, 127(1), pp. 499-508.
Palmer, J.A. (2006) 'Variational and scale mixture representations of non-Gaussian densities for estimation in the Bayesian linear model: Sparse coding, independent component analysis, and minimum entropy segmentation'.
Palmer, J.A., Kreutz-Delgado, K. and Makeig, S. (2006) *International Conference on Independent Component Analysis and Signal Separation*. Springer.
Palmer, J.A., Kreutz-Delgado, K., Rao, B.D. and Makeig, S. (2007) *International Conference on Independent Component Analysis and Signal Separation*. Springer.
Pettersson, K., Jagadeesan, S., Lukander, K., Henelius, A., Haeggstrom, E. and Muller, K. (2013a) 'Algorithm for automatic analysis of electro-oculographic data', *Biomed Eng Online*, 12, p. 110.
Pettersson, K., Jagadeesan, S., Lukander, K., Henelius, A., Hæggström, E. and Müller, K. (2013b) 'Algorithm for automatic analysis of electro-oculographic data', *Biomedical engineering online*, 12(1), p. 1.
Pham, T.T., Croft, R.J., Cadusch, P.J. and Barry, R.J. (2011) 'A test of four EOG correction methods using an improved validation technique', *Int J Psychophysiol*, 79(2), pp. 203-10.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson, R., Miller, G.A., Ritter, W., Ruchkin, D.S. and Rugg, M.D. (2000) 'Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria', *Psychophysiology*, 37(02), pp. 127-152.
Picton, T.W., Goodman, W.S. and Bryce, D.P. (1970) 'Amplitude of evoked responses to tones of high intensity', *Acta Otolaryngol*, 70(2), pp. 77-82.
Picton, T.W., Lins, O.G. and Scherg, M. (1995) 'The recording and analysis of event-related potentials', *Handbook of neuropsychology*, 10, pp. 3-3.
Plöchl, M., Ossandón, J.P. and König, P. (2012) 'Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data', *Frontiers in human neuroscience*, 6, p. 278.
Ponomarev, V.A., Gurskaia, O.E., Kropotov Iu, D., Artiushkova, L.V. and Muller, A. (2010a) '[The comparison of clustering methods of EEG independent components in healthy subjects and patients with post concussion syndrome after traumatic brain injury]', *Fiziol Cheloveka*, 36(2), pp. 5-14.
Ponomarev, V.A., Gurskaya, O.E., Kropotov, Y.D., Artjushkova, L.V. and Müller, A. (2010b) 'Comparison of methods for clustering independent EEG components in healthy subjects and patients with postconcussion syndrome after traumatic brain injury', *Human Physiology*, 36(2), pp. 123-131.
Romero, S., Mananas, M.A. and Barbanoj, M.J. (2008) 'A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: a simulation case', *Comput Biol Med*, 38(3), pp. 348-60.
Romero, S., Mañanas, M.A. and Barbanoj, M.J. (2009) 'Ocular reduction in EEG signals based on adaptive filtering, regression and blind source separation', *Annals of biomedical engineering*, 37(1), pp. 176-191.
Romo-Vazquez, R., Ranta, R., Louis-Dorr, V. and Maquin, D. (2007) *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE.
Ross, R.M., McNair, N.A., Fairhall, S.L., Clapp, W.C., Hamm, J.P., Teyler, T.J. and Kirk, I.J. (2008) 'Induction of orientation-specific LTP-like changes in human visual evoked potentials by rapid sensory stimulation', *Brain Res Bull*, 76(1-2), pp. 97-101.
Rousselet, G.A. (2012) 'Does Filtering Preclude Us from Studying ERP Time-Courses?', *Front Psychol*, 3, p. 131.
Sakai, K. and Miyashita, Y. (1991) 'Neural organization for the long-term memory of paired associates', *Nature*, 354(6349), pp. 152-5.

Schroder, A., van Diepen, R., Mazaheri, A., Petropoulos-Petalas, D., Soto de Amesti, V., Vulink, N. and Denys, D. (2014) 'Diminished n1 auditory evoked potentials to oddball stimuli in misophonia patients', *Front Behav Neurosci*, 8, p. 123.
Schweitzer, P.K. and Tepas, D.I. (1974) 'Intensity effects of the auditory evoked brain response to stimulus onset and cessation', *Perception & Psychophysics*, 16(2), pp. 396-400.
Sebat, J., Levy, D.L. and McCarthy, S.E. (2009) 'Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders', *Trends Genet*, 25(12), pp. 528-35.
Silverman, J., Buchsbaum, M. and Henkin, R. (1969) 'Stimulus sensitivity and stimulus intensity control', *Percept Mot Skills*, 28(1), pp. 71-8.
Smallwood, N., Spriggs, M.J., Thompson, C.S., Wu, C.C., Hamm, J.P., Moreau, D. and Kirk, I.J. (2015) 'Influence of physical activity on human sensory long-term potentiation', *Journal of the International Neuropsychological Society*, 21(10), pp. 831-840.
Smith, S.W. (1997) 'The scientist and engineer's guide to digital signal processing'.
Sponheim, S.R., Clementz, B.A., Iacono, W.G. and Beiser, M. (1994) 'Resting EEG in first-episode and chronic schizophrenia', *Psychophysiology*, 31(1), pp. 37-43.
Stevens, J.C. and Guiroa, M. (1964) 'Individual loudness functions', *The Journal of the Acoustical Society of America*, 36, pp. 2210-2213.
Stewart, A.X., Nuthmann, A. and Sanguinetti, G. (2014) 'Single-trial classification of EEG in a visual object task using ICA and machine learning', *Journal of neuroscience methods*, 228, pp. 1-14.
Sutton, S., Braren, M., Zubin, J. and John, E.R. (1965) 'Evoked-Potential Correlates of Stimulus Uncertainty', *Science*, 150(3700), pp. 1187-1188.
Tanner, D., Morgan-Short, K. and Luck, S.J. (2015) 'How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition', *Psychophysiology*, 52(8), pp. 997-1009.
Tepas, D.I., Boxerman, L.A. and Anch, A.M. (1972) 'Auditory evoked brain responses: Intensity functions from bipolar human scalp recordings', *Perception & Psychophysics*, 11(3), pp. 217-221.
Teplan, M. (2002) 'Fundamentals of EEG measurement', *Measurement science review*, 2(2), pp. 1-11.

Teyler, T.J., Hamm, J.P., Clapp, W.C., Johnson, B.W., Corballis, M.C. and Kirk, I.J. (2005) 'Long-term potentiation of human visual evoked responses', *Eur J Neurosci*, 21(7), pp. 2045-50.
Tong, L., Liu, R.W., Soon, V.C. and Huang, Y.F. (1991) 'Indeterminacy and identifiability of blind identification', *IEEE Transactions on circuits and systems*, 38(5), pp. 499-509.
Tsumoto, T. and Suda, K. (1979) 'Cross-depression: an electrophysiological manifestation of binocular competition in the developing visual cortex', *Brain research*, 168(1), pp. 190-194.
Turetsky, B.I., Raz, J. and Fein, G. (1988) 'Noise and signal power and their effects on evoked potential estimation', *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 71(4), pp. 310-318.
Uriguen, J.A. and Garcia-Zapirain, B. (2015) 'EEG artifact removal-state-of-the-art and guidelines', *J Neural Eng*, 12(3), p. 031001.
Vanrullen, R. (2011) 'Four common conceptual fallacies in mapping the time course of recognition', *Front Psychol*, 2, p. 365.
Verleger, R. (1991) 'The instruction to refrain from blinking affects auditory P3 and N1 amplitudes', *Electroencephalography and Clinical Neurophysiology*, 78(3), pp. 240-251.
Verleger, R., Gasser, T. and Mocks, J. (1982) 'Correction of EOG artifacts in event-related potentials of the EEG: aspects of reliability and validity', *Psychophysiology*, 19(4), pp. 472-80.
Vigario, R. and Oja, E. (2008) 'BSS and ICA in neuroinformatics: from current practices to open challenges', *IEEE Rev Biomed Eng*, 1, pp. 50-61.
Vigário, R., Särelä, J. and Oja, E. (1998) 'Independent component analysis in wave decomposition of auditory evoked fields', in *ICANN 98*. Springer, pp. 287-292.
Viola, F.C., Thorne, J.D., Bleeck, S., Eyles, J. and Debener, S. (2011) 'Uncovering auditory evoked potentials from cochlear implant users with independent component analysis', *Psychophysiology*, 48(11), pp. 1470-80.
Vogel, E.K. and Luck, S.J. (2000) 'The visual N1 component as an index of a discrimination process', *Psychophysiology*, 37(2), pp. 190-203.
von Knorring, L., Monakhov, K. and Perris, C. (1978) 'Augmenting/reducing: an adaptive switch mechanism to cope with incoming signals in healthy subjects and psychiatric patients', *Neuropsychobiology*, 4(3), pp. 150-79.

Vossen, H., Van Breukelen, G., Hermens, H., Van Os, J. and Lousberg, R. (2011) 'More potential in statistical analyses of event-related potentials: a mixed regression approach', *Int J Methods Psychiatr Res*, 20(3), pp. e56-68.
Wallstrom, G.L., Kass, R.E., Miller, A., Cohn, J.F. and Fox, N.A. (2004) 'Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods', *International journal of psychophysiology*, 53(2), pp. 105-119.
Walter, W.G. (1964) 'Slow potential waves in the human brain associated with expectancy, attention and decision', *European Archives of Psychiatry and Clinical Neuroscience*, 206(3), pp. 309-322.
Weerts, T.C. and Lang, P.J. (1973) 'The effects of eye fixation and stimulus and response location on the contingent negative variation (CNV)', *Biological Psychology*, 1(1), pp. 1-19.
Whitton, J.L., Lue, F. and Moldofsky, H. (1978) 'A spectral method for removing eye movement artifacts from the EEG', *Electroencephalogr Clin Neurophysiol*, 44(6), pp. 735-41.
Widmann, A. and Schroger, E. (2012) 'Filter effects and filter artifacts in the analysis of electrophysiological data', *Front Psychol*, 3, p. 233.
Widmann, A., Schroger, E. and Maess, B. (2015) 'Digital filter design for electrophysiological data--a practical approach', *J Neurosci Methods*, 250, pp. 34-46.
Winkler, I., Debener, S., Muller, K.R. and Tangermann, M. (2015) 'On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP', *Conf Proc IEEE Eng Med Biol Soc*, 2015, pp. 4101-5.
Woestenburg, J.C., Verbaten, M.N. and Slangen, J.L. (1983) 'The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain', *Biol Psychol*, 16(1-2), pp. 127-47.
Woldorff, M.G. (1993) 'Distortion of ERP averages due to overlap from temporally adjacent ERPs: analysis and correction', *Psychophysiology*, 30(1), pp. 98-119.
Woodman, G.F. (2010) 'A brief introduction to the use of event-related potentials in studies of perception and attention', *Atten Percept Psychophys*, 72(8), pp. 2031-46.
Woodman, G.F. and Luck, S.J. (2003) 'Serial deployment of attention during visual search', *J Exp Psychol Hum Percept Perform*, 29(1), pp. 121-38.
Woods, D.L. (1995) 'The component structure of the N1 wave of the human auditory evoked potential', *Electroencephalogr Clin Neurophysiol Suppl*, 44, pp. 102-9.

Wyss, C., Hitz, K., Hengartner, M.P., Theodoridou, A., Obermann, C., Uhl, I., Roser, P., Grunblatt, E., Seifritz, E., Juckel, G. and Kawohl, W. (2013) 'The loudness dependence of auditory evoked potentials (LDAEP) as an indicator of serotonergic dysfunction in patients with predominant schizophrenic negative symptoms', *PLoS One*, 8(7), p. e68650.

Zeman, P.M., Till, B.C., Livingston, N.J., Tanaka, J.W. and Driessen, P.F. (2007) 'Independent component analysis and clustering improve signal-to-noise ratio for statistical analysis of event-related potentials', *Clin Neurophysiol*, 118(12), pp. 2591-604.

Zhang, F., Hammer, T., Banks, H.L., Benson, C., Xiang, J. and Fu, Q.J. (2011) 'Mismatch negativity and adaptation measures of the late auditory evoked potential in cochlear implant users', *Hear Res*, 275(1-2), pp. 17-29.

Zhang, L.I., Hui-zhong, W.T. and Poo, M.-m. (2000) 'Visual input induces long-term potentiation of developing retinotectal synapses', *Nature neuroscience*, 3(7), pp. 708-715.

Zouridakis, G., Simos, P.G. and Papanicolaou, A.C. (1998) 'Multiple bilaterally asymmetric cortical sources account for the auditory N1m component', *Brain Topogr*, 10(3), pp. 183-9.

Zuckerman, M., Murtaugh, T. and Siegel, J. (1974) 'Sensation seeking and cortical augmenting-reducing', *Psychophysiology*, 11(5), pp. 535-42.