# APPLYING COGNITIVE ELECTROPHYSIOLOGY TO NEURAL MODELLING OF THE ATTENTIONAL BLINK

a thesis submitted to

The University of Kent at Canterbury

in the subject of computer science

for the degree

of doctor of philosophy

By

Patrick Craston

July 2008

# Contents

**List of tables** **ix**

**List of figures** **xxi**

**Abstract** **xxii**

**Acknowledgements** **xxiii**

**I Introduction** **1**

**1 Introduction** **2**

1.1 Overview . . . . . 2

1.1.1 Why do we need to study temporal attention? . . . . . 3

1.1.2 Experimental investigations using cognitive electrophysiology . . . . 3

1.1.3 Modelling cognition . . . . . 4

1.2 Motivation & contribution . . . . . 5

1.2.1 Investigating temporal attention in humans . . . . . 6

1.2.2 Applying cognitive electrophysiology to neural modelling . . . . . 9

1.2.3 Using temporal attention research for computer systems design . . . 10

1.3 Publications . . . . . 11

1.3.1 Personal contribution to the collaborative research . . . . . 11

1.4 Organisation . . . . . 12

1.4.1 Part I . . . . . 12

1.4.2 Part II . . . . . 12

1.4.3 Part III . . . 12
1.4.4 Part IV . . . 13

**2 Literature review** **14**
2.1 Electroencephalography . . . 14
2.1.1 Event-related potentials . . . 15
2.1.2 ERSP and ITC time-frequency analysis . . . 19
2.1.3 EEG source localisation . . . 21
2.2 Computational modelling of cognition . . . 22
2.2.1 Symbolic models . . . 23
2.2.2 Subsymbolic models . . . 23
2.3 Visual processing under high temporal demands . . . 25
2.3.1 Rapid serial visual presentation . . . 25
2.3.2 The attentional blink . . . 26

**II Applying cognitive electrophysiology to neural modelling of the attentional blink** **35**

**3 Theories of the attentional blink** **36**
3.1 Informal theories of the attentional blink . . . 36
3.1.1 Distractor induced suppression . . . 36
3.1.2 Two-stage theory . . . 37
3.1.3 Interference theory and resource sharing . . . 38
3.1.4 Temporary loss of control . . . 40
3.2 Formal models of the attentional blink . . . 40
3.2.1 Global workspace model . . . 41
3.2.2 The CODAM model . . . 43
3.2.3 The LC-NE model . . . 45
3.2.4 The boost and bounce model . . . 47
3.3 The $ST^2$ model . . . 50
3.3.1 Types & tokens . . . 51

3.3.2 Model architecture . . . 51
3.3.3 How the ST$^2$ model blinks . . . 53
3.3.4 Episodic ST$^2$ model . . . 54
3.3.5 Neural activation patterns from the ST$^2$ model . . . 56
3.4 Summary . . . 57

**4 Virtual ERPs from the ST$^2$ model 58**
4.1 Motivation . . . 58
4.2 Hypothesis . . . 59
4.3 A methodology for virtual ERPs . . . 61
4.3.1 Node potentials in the ST$^2$ model . . . 61
4.3.2 Neural correlates of human ERPs . . . 62
4.3.3 Choosing a neural network node potential . . . 62
4.3.4 Virtual ERP averaging procedure . . . 63
4.3.5 Virtual grand average ERPs and virtual ERPimages . . . 63
4.3.6 A word of caution . . . 63
4.4 Changes to the ST$^2$ model in comparison to Bowman and Wyble (2007) . . . 64
4.5 Generating virtual ERP components . . . 64
4.5.1 Early visual processing . . . 65
4.5.2 Working memory encoding . . . 65
4.5.3 Attentional selection . . . 67
4.6 Summary . . . 68

**III Using virtual and human ERPs to explore the limits of conscious perception 69**

**5 How distractors influence target selection in RSVP 70**
5.1 Introduction . . . 71
5.1.1 RSVP without distractors: Skeletal presentation . . . 71
5.1.2 Reducing backward masking in RSVP . . . 74
5.1.3 Motivation and overview . . . 74

5.2 Methods . . . 75
5.2.1 EEG methods . . . 75
5.3 Target processing in skeletal presentation . . . 75
5.3.1 Results: Behaviour . . . 75
5.3.2 Results: Electrophysiology . . . 76
5.3.3 Modelling skeletal presentation . . . 78
5.3.4 Discussion . . . 84
5.4 Removing the T+1 distractor . . . 91
5.4.1 Results . . . 91
5.4.2 Discussion . . . 93
5.5 Conclusion . . . 94

**6 The attentional blink reveals serial working memory encoding 96**
6.1 Introduction . . . 97
6.1.1 Resource sharing vs. two-stage theories . . . 98
6.1.2 The P3 component as a measure of resource allocation? . . . 98
6.1.3 Overview . . . 100
6.2 Methods . . . 101
6.2.1 Experiment 1 . . . 101
6.2.2 Experiment 2 . . . 101
6.3 Results: Experiment 1 . . . 102
6.3.1 Behaviour . . . 102
6.3.2 Human ERP . . . 103
6.3.3 Virtual ERP . . . 103
6.4 Results: Experiment 2 . . . 106
6.4.1 Behaviour . . . 106
6.4.2 Human ERPs . . . 107
6.4.3 Virtual ERPs . . . 109
6.5 Discussion . . . 111
6.5.1 The meaning of P3 amplitude for targets in RSVP . . . 111
6.5.2 Working memory encoding is serial during the attentional blink . . . 112

6.5.3 Interference between T1 and T2 at lag 1 . . . 113
6.5.4 Evaluating previous findings . . . 114
6.6 Conclusion . . . 120

**7 Temporal variation in target processing during the attentional blink 121**
7.1 Introduction . . . 122
7.1.1 Transient attention and the $ST^2$ model . . . 122
7.1.2 Increased temporal variance in target processing during the AB? . . . 123
7.1.3 Overview . . . 124
7.2 Methods . . . 125
7.2.1 EEG analysis . . . 125
7.3 Results & discussion . . . 127
7.3.1 Behavioural results . . . 127
7.3.2 Evidence for increased variance during the AB . . . 127
7.3.3 Verifying our analysis of phase distributions . . . 131
7.3.4 Virtual ERPs from the $ST^2$ model . . . 136
7.4 Conclusion . . . 140

**8 The attentional blink modulates the influence of target strength on conscious perception 142**
8.1 Introduction . . . 143
8.1.1 The influence of bottom-up strength and attention on perception . . . 143
8.1.2 The P3 as a correlate of conscious perception . . . 144
8.1.3 Overview . . . 144
8.2 Methods . . . 145
8.3 Results . . . 146
8.3.1 Behavioural . . . 146
8.3.2 EEG . . . 147
8.4 Discussion . . . 151
8.4.1 Target difficulty affects the P3 for targets outside but not inside the AB . . . 151
8.4.2 Virtual ERPs from the $ST^2$ model . . . 152

8.4.3 The two-phase strength sensitivity theory . . . . . 155
8.5 Conclusion . . . . . 169

**IV Further discussion & conclusion** **171**

**9 Using EEG to design adaptive computer systems** **172**
9.1 Introduction . . . . . 172
9.2 Method . . . . . 173
9.3 Summary of results . . . . . 174
9.3.1 Separate P3 and alpha analysis . . . . . 175
9.3.2 Combined P3 and alpha analysis . . . . . 175
9.4 The practical details of such a device . . . . . 176
9.5 Evaluating the approach . . . . . 178

**10 Summary & conclusion** **179**
10.1 Summary . . . . . 179
10.2 Contributions . . . . . 181
10.2.1 The influence of distractors on target processing . . . . . 181
10.2.2 The meaning of P3 amplitude in RSVP . . . . . 182
10.2.3 Implications for theories of the attentional blink . . . . . 183
10.2.4 Virtual ERPs as an additional dimension of neural modelling . . . . 186
10.2.5 Virtual ERPs assist electrophysiological experimentation . . . . . 187
10.3 Future work . . . . . 187
10.3.1 Additional evidence elucidating the nature of working memory encoding during the AB . . . . . 187
10.3.2 Dissecting the virtual ERP to identify the neural substrates of the human EEG . . . . . 189
10.3.3 'Lesioning' the $ST^2$ model . . . . . 190
10.3.4 Virtual fMRI traces? . . . . . 191
10.3.5 Modelling blinkers and non-blinkers . . . . . 192

**A Methods** **195**

A.1 Experiment 1 . . . . . 195

A.1.1 Participants . . . . . 195

A.1.2 Stimuli and apparatus . . . . . 195

A.1.3 Procedure . . . . . 195

A.1.4 EEG recording . . . . . 197

A.1.5 EEG data analysis . . . . . 198

A.1.6 Computational modelling . . . . . 198

A.2 Experiment 2 . . . . . 199

A.2.1 Participants . . . . . 199

A.2.2 Stimuli and apparatus . . . . . 199

A.2.3 EEG recording . . . . . 201

A.2.4 EEG data analysis . . . . . 201

A.2.5 Computational modelling . . . . . 202

**Bibliography** **203**

# List of tables

1 List of weights that were modified during this work. The values from Bowman & Wyble, 2007 are shown in brackets. All other parameters remained unchanged. . . . . 64
2 Results from the combined P3 and alpha detection algorithm. . . . . 175

# List of figures

1 Averaging segments of raw EEG to extract event-related potentials. Positive is plotted upwards. . . . . 15

2 Panel A: A sample ERP showing the P1, N1 and P3 ERP components. Positive is plotted upwards. Panel B: A sample ssVEP wave oscillating at 10Hz. Positive is plotted upwards. Panel C: An illustration of how to extract a lateralised ERP component, such as the N2pc. Figure adapted from Woodman and Luck (2003). In this figure negative is plotted upwards. . . . . 16

3 Panel A: One-dimensional ERP plot containing EEG activity averaged over time. Positive is plotted upwards. Panel B: ERPimage plot: Time is plotted on the x-axis and individual EEG epochs are plotted on the y-axis. Voltage values are displayed using a colour coded scale and the data is smoothed along the y-axis to enhance the visual signal-to-noise ratio. . . . . 18

4 Panel A: Event-related spectral perturbation (ERSP) plot. Time is plotted on the x-axis and the y-axis contains the frequency bands of the specified range. Power is displayed using the colour scales specified to the right of the plot. The waveform to the left shows the baseline average power per frequency band. The waveform below depicts the highest and lowest mean power values across all frequencies relative to baseline at each time point. Panel B: Inter-trial coherence (ITC) plot. Time is plotted on the x-axis and the y-axis contains the frequency bands of the specified range. ITC phase coherence values are displayed using the colour scales specified to the right. The plot to the left depicts the minimum and maximum mean ITC values per frequency across all time points. The waveform below shows the grand average ERP waveform. . . . . 19

5 Example of a simple neural network containing three layers. In such a neural network, the input layer will typically receive an input signal in the form of some activation pattern. This activation feeds through the hidden layer where it is modulated according to the weight values of the connections between the layers. The output layer then provides the user with an activation pattern that can be interpreted. . . . . 24

6 The RSVP paradigm that is typically used in attentional blink experiments. Panel A: A letters-in-digits attentional blink task as used in Chun and Potter (1995). Panel B: A colour-marked attentional blink task as used in Raymond, Shapiro, and Arnell (1992) . . . . . 26

7 The attentional blink. T2 accuracy is conditional on correct report of T1 (T2|T1), raw T1 accuracy and percentage of swaps. The data is from Chun and Potter (1995). . . . . 27

8 Accuracy per target position for (a) T1 T2 T3 T4: No AB if four targets are presented in a row. (b) T1 D T2 T3: If T1 is followed by a distractor, then T2 is 'blinked', however, T3 is 'spared'. (c) T1 T2 D T3: T2 is presented following T1 and is 'spared', T2 and T3 are separated by a distractor and T3 is 'blinked'. The data from curve (a) is from Nieuwenstein and Potter (2006). The data in curves (b) and (c) are from Olivers, van der Stigchel, and Hulleman (2007). . . . . 29

9 The AB is reduced if the distractors following each of the targets (T1 and T2) are removed. T2 accuracy conditional on correct report of T1 (T2|T1), T2|T1 accuracy if T1 is followed by a blank (T1+1) reproduced from Chun and Potter (1995). T2|T1 accuracy if T2 is placed at the end of the RSVP stream reproduced from Giesbrecht and Di Lollo (1998). . . . . 30

10 Panel A: Response distribution (percentage of trials per visibility score) for a T2 presented during the AB, i.e. T2 follows T1 at lag 3 and participants are instructed to report both T1 and T2. Adapted from Sergent and Dehaene (2004). Panel B: Response distribution (percentage of trials per visibility score) for a T2 presented outside the AB, i.e. T2 follows T1 at lag 3 but participants are instructed to ignore T1. Adapted from Sergent, Baillet, and Dehaene (2005). . . . . 32

11 Panel A: ERPimage plot of the P3 components evoked by T1 and T2 when T2 is presented during the AB (at lag 3). Trials are sorted by visibility score, lowest visibility at the bottom. Voltage values (positive in red and negative in blue) are smoothed over 50 trial windows. Panel B: Histogram of mean amplitude of the T2 P3 component (presented during the AB) per visibility category. Categories 4 and 3 correspond to visibility scores $> 50\%$ ('seen' trials), whereas categories 2 and 1 to visibility scores $< 50\%$ ('unseen' trials). Within seen and unseen trials, the categories are classified per participant using the median of that participant's response distribution. Adapted from Sergent et al. (2005). . . . . 32

12 Neuroanatomical substrates of the AB. The gold circles represent activation foci from fMRI studies, whereas the light- and dark-green shaded regions represent areas implicated in lesion and MEG studies, respectively. Adapted from Marois (2005). . . . . 33

13 Panel A: Processing pathways in the global workspace model. Panel B: Areas C and D represent conscious perception, hence, the corresponding curves of the figure reflect the simulated behavioural accuracy of the global workspace model during the AB. Panel C: Neural activity evoked by seen and missed targets during the AB. Areas A2 and B2 correspond to early sensory processing, whereas areas C and D correspond to later stages of processing. Adapted from Dehaene, Sergent, and Changeux (2003). . . . . 41

14 The version of the CODAM model used to simulate the attentional blink. Adapted from Fragopanagos, Kockelkoren, and Taylor (2005). . . . . 43

15 Panel A: Membrane Potentials from CODAM for T1. Panel B: Membrane Potentials from CODAM for T2 presented during the AB. Adapted from Fragopanagos et al. (2005). . . . . 44

16 Depiction of the locus coeruleus - norepinephrine (LC-NE) model. Adapted from Nieuwenhuis, Gilzenrat, Holmes, and Cohen (2005). . . . . 46

17 Activation dynamics of the abstracted locus coeruleus (LC) for missed and seen T2s during the AB. LC activity and norepinephrine (NE) output are scaled on separate axes. Time 0 ms indicates the onset of the simulated trials. Adapted from Nieuwenhuis, Gilzenrat, et al. (2005). . . . . 47

18 Schematic depiction of the boost and bounce model. As stimuli representations feed through the model, they are processed for colour, shape and identity. The task the observer is asked to do (report the letter among the digits in the example shown here) determines the attentional set, which is implemented in the working memory gating system. The gating system works by selectively enhancing ('boosting') targets and inhibiting ('bouncing') distractors and thus allowing targets to enter working memory. Adapted from Olivers and Meeter (2008). . . . . 48

19 Neural dynamics of the boost and bounce model for a target presented during the AB (T2 presented at lag 2). Panel A: Activation from nodes resembling bottom-up sensory processing. Panel B: Top-down attentional response to the input (combination of both excitatory and inhibitory feedback). Panel C: Combined activation trace of bottom-up and top-down signal. Adapted from Olivers and Meeter (2008). . . . . 49

20 The $ST^2$ model: (1) Input & extraction of types in stage one, (2) Working memory tokens in stage two, (3) Temporal attention from the blaster. Adapted from Bowman and Wyble (2007). . . . . 52

21 Panel A: Human accuracy: Basic AB (T2 accuracy conditional on T1 correct, T2|T1), raw T1 accuracy, Swaps. Reproduced from Chun and Potter (1995). Panel B: $ST^2$ model accuracy: Basic AB (T2 accuracy conditional on T1 correct), raw T1 accuracy, swaps. . . . . 54

22 Panel A: Competitive regulation of attention; bottom-up target input attempts to trigger (excites) attention while the working memory encoding process attempts to shut off (suppresses) attention. Panel B upper figure: Activation dynamics of the $eST^2$ model when multiple targets are presented in a row (i.e. whole report). Panel B lower figure: Activation dynamics of the $eST^2$ model during the AB. Adapted from Wyble, Bowman, and Nieuwenstein (2009). . . . . 55

23 The hypothesis underlying the virtual ERP approach. . . . . 60

24 A typical pair of neural network nodes situated in two neighbouring connected layers of the $ST^2$ model. As shown in the figure, we can extract the membrane potential, presynaptic activation and postsynaptic activation for each node of the $ST^2$ model's neural network implementation (neural-$ST^2$). 61

25 Panel A: Virtual SSVEP component from input and masking layers resembling early visual processing. Panel B: Virtual P3 component from item layer, TFL, binding pool gate nodes and token gate nodes, resembling working memory encoding. Panel C: Virtual N2pc component from the blaster node reflecting the firing of attentional enhancement, which initiates attentional selection. In all figures, time point 0 corresponds to the onset of the RSVP stream. . . . . 66

26 Panel A: A regular RSVP stream where a target letter is embedded in a stream of digit distractors. Panel B: The skeletal presentation paradigm, which contains only the target letter and the following digit distractor as its mask. . . . . 71

27 Behavioural accuracy scores from AB studies using the skeletal paradigm. Panel A: White circles (T1) show T1 accuracy per lag. Black squares (T2|T1 correct) indicate T2 accuracy per lag conditional on T1 being correct. Adapted from McLaughlin, Shore, and Klein (2001). Panel B: White circles (Single task) show T2 accuracy per lag when subjects were instructed to ignore T1 and report T2. Black circles (Dual task) show T2 accuracy per lag when subjects were instructed to report both T1 and T2 per lag. Adapted from Rolke, Bausenhart, and Ulrich (2007). . . . . 73

28 Panel A: Human ERP for targets in RSVP and skeletal presentation averaged across P7 and P8 electrode locations. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target. Positive is plotted upwards. Panel B: Fast fourier transform (FFT) of the ERP from the P7 and P8 electrode locations for the RSVP (left) and skeletal (right) condition. The RSVP condition shows a peak in the FFT plot at the frequency of target presentation (approx. 21Hz), which is not present for skeletal presentation. . . . . 77

29 A virtual ssVEP wave for the RSVP and vERP early components for the skeletal condition from input and masking layer of the $ST^2$ model. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target. . . . . 79

30 Step 2 of simulating skeletal presentation. As indicated in the figure, the connection from stage one that triggers the blaster is moved from the task filtered layer to the masking layer. . . . . 80
31 After Step 2: Virtual P3 for the RSVP and skeletal condition. The RSVP vERP is baseline corrected to -200 to 0ms with respect to target onset to account for distractor related activity, which is absent in the skeletal RSVP condition. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target. . . . . 81
32 Step 3 of simulating skeletal presentation. As indicated in the figure, the weight of the connection from task demand to target nodes in the TFL is reduced by 0.92%. . . . . 83
33 After Step 3: Virtual P3 for the RSVP and skeletal condition. The RSVP vERP is baseline corrected to -200 to 0ms with respect to target onset to account for distractor related activity, which is absent in the skeletal RSVP condition. 'T' indicates the presentation of the target and ERPs time locked to presentation of the target. . . . . 84
34 Virtual N2pc component from the $ST^2$ model's blaster output nodes for RSVP and skeletal presentation. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target. . . . . 88
35 Panel A: Human ERP for targets in the RSVP and T+1 blank conditions averaged across P7 and P8 electrode locations. Positive is plotted upwards. Panel B: Virtual P3 from the $ST^2$ model for the RSVP and T+1 blank conditions. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target. . . . . 92
36 Panel A: hERP P3 component from Pz for the easy and hard condition. Positive is plotted upwards. Panel B: $ST^2$'s vERP containing the virtual P3 component for the easy and hard condition. For both panels, 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target. . . . . 104

37 Panel A: Human behavioural accuracy data for lag 1, lag 3 and lag 8. Panel B: Simulated behavioural accuracy of the $ST^2$ model for lag 1, lag 3 and lag 8. Circles indicate T2 accuracy conditional on correct T1 report, triangles represent raw T1 accuracy and squares indicate swaps, i.e. the condition when T1 and T2 were correctly identified but reported in the wrong order. In panel A, error bars represent standard error of the mean. . . . . 105

38 Panel A: hERP from Pz for 1) the lag 3 noAB condition (T1 and T2 correctly reported), 2) the lag 3 AB condition (T1 accurately identified but T2 not correctly reported), 3) the lag 8 condition (T1 and T2 correctly reported). Positive is plotted upwards. Panel B: $ST^2$'s vERP containing the virtual P3 for 1) the lag 3 noAB condition, 2) the lag 3 AB condition, 3) the lag 8 condition. In both panels 'T1' and 'T2' indicate the presentation of the T1 and T2 respectively, and ERPs are time locked to T1. . . . . 108

39 Panel A: hERP from Pz for 1) the lag 1 condition (T1 and T2 correctly reported), 2) the lag 8 condition (T1 and T2 correctly reported). Positive is plotted upwards. Panel B: $ST^2$'s vERP containing the virtual P3 component for 1) the lag 1 condition, 2) the lag 8 condition. In both panels 'T1' and 'T2' indicate the presentation of the T1 and T2 respectively, and ERPs are time locked to T1. . . . . 110

40 Panel A: P3 phase-sorted ERPimage for targets outside the AB (T2 at lag 8). Panel B: P3 phase-sorted ERPimage for targets inside the AB (T2 at lag 3). The T2 is presented at time point zero. The dashed line indicates the time point around which the phase sorting is centred, corresponding to the peak of the P3 component. . . . . 129

41 Panel A: N2pc phase-sorted ERPimage for targets outside the AB (T2 at lag 8). Panel B: N2pc phase-sorted ERPimage for targets inside the AB (T2 at lag 3). The T2 is presented at time point zero. The dashed line indicates the time point around which the phase sorting is centred, corresponding to the peak of the N2pc component. . . . . 130

42 Panel A: P3 phase-sorted ERPimage for T1 presented at lag 8. Panel B: N2pc phase-sorted ERPimage for T1 presented at lag 8. The T1 is presented at time point zero. The dashed line indicates the time point around which the phase sorting is centred, corresponding to the peak of the P3 and N2pc component respectively. . . . . 132

43 Panel A: P3 phase-sorted ERPimage for the baseline condition (T1 at lag 8 condition but sorted around the -300ms time point with respect to T1 onset). Panel B: N2pc phase-sorted ERPimage for the baseline condition. The T1 is presented at time point zero. The dashed line indicates the time point around which the phase sorting is centred. . . . . 134

44 Panel A: P3 virtual ERPimage for correctly identified targets outside the AB (T2 at lag 8). Panel B: P3 virtual ERPimage for correctly identified targets inside the AB (T2 at lag 3). The target is presented at time point zero. The solid line indicates the peak of the virtual P3 component in that trial. . . . 137

45 Panel A: N2pc virtual ERPimage for correctly identified targets outside the AB (T2 at lag 8. Panel B: N2pc virtual ERPimage for correctly identified targets inside the AB (T2 at lag 3). The target is presented at time point zero. The solid line indicates the peak of the virtual N2pc component in that trial. . . . . 138

46 Panel A: ERPimage from electrode Pz for a target outside the AB. Colour strips to the left of the plot indicate the accuracy and target difficulty for that particular trial. Trials within each accuracy/target difficulty category are randomly shuffled before vertically smoothing to average over individual subject differences. Panel B: Bar chart displaying the mean P3 size (300-600ms with respect to target onset) for each accuracy-target difficulty combination. The error bars depict the standard error of the mean. . . . . 148

47 Panel A: ERPimage from electrode Pz for targets presented inside the AB (T2 at lag 3 following a correctly identified T1). Colour strips to the left of the plot indicate the accuracy and target difficulty for that particular trial. Trials within each accuracy/target difficulty category are randomly shuffled before vertically smoothing to average over individual subject differences. Panel B: Bar chart displaying the mean P3 size (300-600ms with respect to T2 onset) for each accuracy-target difficulty combination. The error bars depict the standard error of the mean. . . . . 149

48 Virtual P3 from the $ST^2$ model for targets outside the AB (a single target in RSVP). Colour strips to the left of the plots indicate the accuracy and target strength for that particular trial of the simulation. . . . . 153

49 Virtual P3 from the $ST^2$ model for targets inside the AB, i.e. a T2 presented at lag 3 following a correctly reported T1. Colour strips to the left of the plots indicate the accuracy and target strength for that particular trial. . . 154

50 The two-phase theory's hypothesised target activation traces for varying target strengths. The figure describes the profile of activation without enhancement from the blaster. . . . . 156

51 Target activation traces from the two-phase theory including blaster enhancement. When targets are presented individually (or outside the AB), the blaster is available and fires as soon as the target has been detected by the system, i.e. during phase 1. For strong targets, the enhancement from the blaster leads to a steady-state activation level where the amount of activation corresponds to the target's initial strength value. Weak targets do not have sufficient activation strength to fire the blaster and their activation decays back to baseline during phase 1. . . . . 158

52 The two-phase theory's hypothesised target activation traces including blaster enhancement. For a target (T2) that is presented during the AB, the blaster is delayed and does not fire until phase 2, as the blaster is suppressed while the preceding target (T1) is encoded into working memory. Due to the delay, only strong targets can fire the blaster, weak targets decay to baseline during phase 1. Note that the figure only contains the activation trace for the T2; T1 activation is not shown. . . . . 160

53 Panel A: Response distribution (percentage of trials in each visibility category) for a T2 presented during the AB, i.e. T2 follows T1 at lag 3 and participants are instructed to report both T1 and T2. Adapted from Sergent and Dehaene (2004). Panel B: Response distribution (percentage of trials in each visibility category) for a T2 presented outside the AB, i.e. T2 follows T1 at lag 3 but participants are instructed to ignore T1. Adapted from Sergent et al. (2005). . . . . 162

54 Panel A: ERPimage plot of the P3 components evoked by T1 and T2 when T2 is presented during the AB (at lag 3). Trials are sorted by visibility score, lowest visibility at the bottom. Voltage values (positive in red and negative in blue) are smoothed over 50 trial windows. Panel B: Histogram of mean amplitude of the T2 P3 component (presented during the AB) per visibility category. Categories 4 and 3 correspond to visibility scores $> 50\%$ ('seen' trials), whereas categories 2 and 1 to visibility scores $< 50\%$ ('unseen' trials). Within seen and unseen trials, the categories are classified per participant using the median of that participant's response distribution. Adapted from Sergent et al. (2005). . . . . 163

55 ERPimage from electrode Pz for a T2 following T1 at lag 1. Colour strips to the left of the plot indicate the accuracy and target difficulty for that particular trial. Trials within each accuracy/target difficulty category are randomly shuffled before vertically smoothing to average over individual subject differences. . . . . 165

56 Bar chart displaying the mean P3 size for each accuracy-target difficulty combination. The error bars depict the standard error of the mean. Panel A: Mean P3 size for the 200-600ms window with respect to T2 onset, which captures the whole P3 component. Panel B: Mean P3 size for the 400-600ms window. This window captures the later part of the P3 component, which is hypothesised to be influenced more strongly by T2 than by T1. . . . . 166

57 Panel A: Illustration of the proposed BBRA system. The probe input provides the time lock signal indicating target presentation. The 'hit/miss' output informs the computer system whether the target was seen by the user or not. Panel B: Schematic diagram showing the electrodes and microcontroller circuit for the BBRA. . . . . 177

58 The experimental paradigm used in Experiment 1. Panel A: An RSVP stream using digits as distractors and a letter as the target. Panel B: A T+1 blank stream where the distractor following the target is omitted. Panel C: A skeletal stream containing only the target letter and the following digit distractor as its mask. . . . . 196

59 The two-target bilateral RSVP paradigm used in Experiment 2. . . . . 200

# Abstract

This thesis proposes a connection between computational modelling of cognition and cognitive electrophysiology. We extend a previously published neural network model of working memory and temporal attention (*Simultaneous Type Serial Token ($ST^2$) model*; Bowman & Wyble, 2007) that was designed to simulate human behaviour during the *attentional blink*, an experimental finding that seems to illustrate the temporal limits of conscious perception in humans. Due to its neural architecture, we can utilise the $ST^2$ model's functionality to produce so-called virtual event-related potentials *(virtual ERPs)* by averaging over activation profiles of nodes in the network. Unlike predictions from textual models, the virtual ERPs from the $ST^2$ model allow us to construe formal predictions concerning the EEG signal and associated cognitive processes in the human brain.

The virtual ERPs are used to make predictions and propose explanations for the results of two experimental studies during which we recorded the EEG signal from the scalp of human participants. Using various analysis techniques, we investigate how target items are processed by the brain depending on whether they are presented individually or during the attentional blink. Particular emphasis is on the P3 component, which is commonly regarded as an EEG correlate of encoding items into working memory and thus seems to reflect conscious perception. Our findings are interpreted to validate the $ST^2$ model and competing theories of the attentional blink. Virtual ERPs also allow us to make predictions for future experiments. Hence, we show how virtual ERPs from the $ST^2$ model provide a powerful tool for both experimental design and the validation of cognitive models.

# Acknowledgements

Most of all, I thank my parents Andrew Craston and Herma Wöbbeking and the rest of my family for their support. I am indebted to Wiebke Berg for her optimism and continuous motivation. This thesis was fuelled by coffee from the best machine in the world!

I wish to thank Howard Bowman for his excellent supervision. Furthermore, I thank Brad Wyble, especially for his patience when working through the model and presenter code, and Srivas Chennu for those inspiring discussions. The experimental work would have not been possible without numerous EEG participants, in particular, Wiebke Berg, Damian Dimmich and Poul Henriksen. Dirk Janssen, Lea Hald and Ingmar Gutberlet were a great help while setting up the EEG experiments. I thank Karl Friston for suggesting to phase sort the ERPimages and Ana Cavalcanti, who introduced me to the idea of doing a PhD in the first place. Furthermore, I am grateful to Ulf Berg and Kristina Berg-Meyer for providing me with accommodation during the write-up. Finally, this PhD was financially supported by a Doctoral Training Account from the UK Engineering and Physical Research Council (EPSRC), the Computing Laboratory at the University of Kent and an EPSRC grant (number GR/S15075/01) awarded to Howard Bowman.

In this spirit: 'We choose to go to the moon and do the other things [such as a PhD covering computational modelling and EEG when coming from a BSc in Computing and Business] ... not because they are easy but because they are hard...' (J. F. Kennedy, 1962)

# Part I

# Introduction

# Chapter 1

# Introduction

## 1.1 Overview

The world we humans live in contains vast amounts of information. While navigating in this world, our sensory organs are subject to a continuous stream of incoming stimuli from various sensory dimensions, such as the visual, auditory or tactile domain. If we were to store and process every detail of this input, this would quickly lead to information-overload. Hence, humans have developed an efficient filtering mechanism, which is commonly referred to as attention (James, 1890; Pashler, 1996).

Attention selects relevant pieces of information at the cost of disregarding others that have been deemed irrelevant by the system (Broadbent, 1958; Deutsch & Deutsch, 1963). The attentional mechanism thus plays a primary role in determining if stimuli are able to enter consciousness (Posner, 1994). In this thesis, we investigate the human attentional system to gain insights into the limits of consciousness when perceiving stimuli from the visual domain.

Research into visual attention can be separated into two areas, namely spatial attention on the one hand, and temporal attention on the other. Whereas the former has been subject to extensive research over the past few decades (e.g. Treisman & Gelade, 1980; Pashler & Badgio, 1985; Chun & Wolfe, 2001), the latter has been studied to a lesser extent in relative terms. This is partly due to the fact that the mechanisms for selecting stimuli in space are much more prominent in our daily lives. Our visual field is constantly filled with multiple

objects and the system is forced to filter out irrelevant items (Wolfe, 1998). The limits of temporal attention, however, are only seldom encountered in our natural environment, as the visual system does remarkably well at processing multiple items occurring within very short time spans (see e.g. Potter, 1976). In fact, multiple stimuli have to appear within less than a second for the visual system to approach its capacity limitations of processing stimuli in time (Lawrence, 1971; Broadbent & Broadbent, 1987; Weichselgartner & Sperling, 1987; Reeves & Sperling, 1986).

### 1.1.1 Why do we need to study temporal attention?

Modern technology has led to a dramatic increase in the pace of our environment. Signals displayed on computer screens quite commonly approach and even exceed the temporal limits of conscious perception. Hence, nowadays the boundaries of temporal attention are salient in daily life and, consequently, the knowledge of these limitations is an important factor when designing modern computer systems (see Chapter 9 for a discussion of this issue).

In the laboratory, researchers use artificial paradigms to study the limitations of temporal attention. One finding, which has received a lot of attention from the scientific community in recent years, is the attentional blink (Raymond et al. (1992); Chun and Potter (1995); see Section 2.3.2 for a detailed description). It describes the observation that, if a stimulus follows another item within approx. half a second in the same spatial location, observers very often miss that second item. If the stimulus is presented on its own, however, it can easily be detected. Consequently, the attentional blink seems to outline the *temporal capacity limitation* of the visual system.

### 1.1.2 Experimental investigations using cognitive electrophysiology

The requirements for conscious perception, as illustrated by target processing during the attentional blink, can be studied using various experimental techniques. Behavioural experiments, which manipulate experimental paradigms and explore the implications on accuracy scores and reaction times, remain the most widely used method of investigation. However,

an increasing number of studies also record neural brain activity while participants are performing a task. When studying temporal attention, one requires a high temporal resolution. This makes electroencephalography (EEG), a non-invasive technique of recording the electrophysiological signal from electrodes placed on the participant's scalp (see Section 2.1), a popular choice for studying temporal attention.

#### Analysing electrophysiological data

A common method for analysing EEG data is the so-called event-related potential technique (ERP; see Section 2.1.1). Hereby, one averages across multiple instances of EEG data time locked to an event of interest. The averaging procedure increases the signal-to-noise ratio by removing EEG activity that is not time locked to the event.

Recently, however, the ERP technique has been criticised for making use of only a subset of the full spectrum of data inherent to the EEG signal (Makeig, Debener, Onton, & Delorme, 2004; Makeig, Delorme, et al., 2004). Accordingly, a cognitive event can produce different kinds of neural responses. For instance, the cognitive event might cause an increase of the number of neurons that are firing or it might modulate the extent to which the neurons fire in synchrony. These processes can have different effects on the EEG. From the grand average ERP, however, one cannot distinguish between these underlying processes (see Chapter 7 for a discussion of this issue). Furthermore, the ERP does not provide information about single trial dynamics. Alternative techniques, which make use of an 'information-based approach to modelling electroencephalographic (EEG) dynamics' (Makeig, Debener, et al., 2004), have thus been proposed. Amongst these, ERPimages (see Section 2.1.1) illustrate trial-by-trial fluctuations of the raw EEG underlying the ERP, whereas inter-trial coherence plots (see Section 2.1.2) display the amount of phase locking present in the EEG data.

### 1.1.3 Modelling cognition

The data-driven approaches to the investigation of cognition described in the previous paragraph, however, do not remove the need for theory. Rather, the large amounts of information inherent to the EEG signal as well as the vast number of possible analyses of the data make it even more important to have distinctive a priori hypotheses (Picton et

al., 2000). A cognitive model provides a unified framework, which can be used to make predictions about the results of a given experimental manipulation.

One approach to modelling uses box-and-arrow diagrams. Although box-and-arrow models can help express theoretical concepts, the predictions they generate are of an informal nature and, thus, can be susceptible to inconsistent interpretation (see Section 2.2 and also Rabbitt, 1993).

Hence, there is a need for models that employ computational methods, which represent the theory in a quantitative manner and provide a platform for formally validating the theoretical hypothesis (R. Cooper, Fox, Farringdon, & Shallice, 1996; see Section 2.2). The benefits of a computational model can be expressed in the following terms. First, a computational model provides directly verifiable predictions about experimental data. The more concrete - and hence unambiguous - these predictions are, the better they can benefit the design of experiments in order to validate these predictions. Second, the computational model provides a formal platform for assessing the theory underlying the model. In designing a computational model, one has to formally commit oneself to certain theoretical standpoints.

Theories expressed in box-and-arrow diagrams can sometimes risk becoming *unfalsifiable*. A theory expressed by means of a computational model, however, can be verified (and consequently also disproved) by experimental data, which accelerates the progress of scientific research (see also Popper, 1959).

## 1.2 Motivation & contribution

This thesis provides a twofold contribution to scientific knowledge. In the following, we introduce the theoretical and methodological contributions of this thesis. First, we present a number of experimental results (of both behavioural and electrophysiological nature) that inform current theories of the temporal attentional system in humans. These findings and their implications are briefly introduced in Section 1.2.1 and are presented in full detail in Chapters 5, 6, 7 and 8. Second, we present a novel technique of extending a previously published cognitive neural network model (Bowman & Wyble, 2007) to simulate 'artificial'

EEG traces, so-called *virtual ERPs*. The hypothesis underlying this methodological contribution is briefly introduced in Section 1.2.2 and then extensively described in Chapter 4. Finally, we describe a potential practical application of this work for the design of computer systems, which is further discussed in Chapter 9.

### 1.2.1 Investigating temporal attention in humans

The experimental data presented in this thesis are used to test a number of hypotheses concerning the mechanisms underlying temporal attention in humans. We briefly introduce these theoretical contributions grouped by their theoretical similarities in the following section.

#### Rapid serial visual presentation enforces late selection

The context of presentation is highly relevant to the way a target item is processed by the visual system. If a target item is presented individually, the system usually has little trouble detecting that target and consolidating it into working memory. Using special experimental paradigms, such as rapid serial visual presentation (RSVP), however, target representations can be rendered sufficiently weak so that the visual system has trouble detecting or sometimes even completely fails to perceive them (see Section 2.3.1). One of these paradigms is the *category-distinguished* RSVP task, where the target does not differ in terms of its visual features but can only be distinguished from distractors by its categorical properties (for instance when the target is a letter among digit distractors).

In Chapter 5, we present experimental results that investigate how targets are processed by the visual system depending on whether they are presented individually (using the so-called skeletal task) or in an RSVP paradigm. We analyse both participants' behavioural responses and also their EEG signal in the form of the event-related potential (ERP) wave (see Section 2.1.1) evoked by targets that are presented either in skeletal presentation or in RSVP. The behavioural results suggest that observers perform worse at detecting targets presented in RSVP. Based on our EEG results, we hypothesise that because RSVP targets are embedded into a continuous stream of distractors, this enforces a late selection strategy. Unlike skeletal presentation where a target is marked by visual onset, a category-distinguished RSVP task requires the system to process all items to a relatively high level

of detail before it can separate targets from distractors. This hypothesis is illustrated by means of virtual ERPs from the $ST^2$ model (see Chapter 4 for a description of virtual ERP methodology and Section 3.3 for a description of the $ST^2$ model) and supported by the human EEG data.

**Attentional selection and working memory encoding during the attentional blink**

As previously discussed, the attentional blink (AB) seems to provide insights into the temporal limits of the human attentional system. The AB describes the finding that people often fail to detect a second target if it appears within approx. half a second of an identified first target. However, if two targets are presented in immediate succession, performance at detecting both targets is excellent (see Section 2.3.2 for an elaborate description of the AB). The literature contains a number of competing theories of the AB (see Chapter 3 for a review), which propose various explanations for the mechanisms underlying temporal attention in humans. In this thesis, we investigate the AB by analysing the EEG correlates of attentional selection and working memory encoding. These results are used to assess the $ST^2$ model against its competitor theories of the AB.

Chapter 6 investigates how targets are encoded into working memory during the AB. Our EEG results suggest that if targets are presented in immediate succession (and within a short time period of less than approx. 150ms), they are encoded into working memory together. Using the $ST^2$ model, we hypothesise that under these circumstances both targets are perceived in a single episode. However, if the targets are at least 200ms apart, the EEG results suggest that working memory encoding occurs in a serial fashion and the targets are encoded separately. If the second target is presented within 200-600ms, it occurs while the first episode is underway, which results in impaired accuracy at detecting the second target and this finding is what is commonly referred to as the AB. Our EEG results are in contrast with pure competition based accounts of the AB, but support theories - such as the $ST^2$ model - that propose a serial nature of working memory during the AB.

In Chapter 7, we further investigate the influence of the AB on how the visual system processes target items. Previous research has suggested increased temporal variance in the processing of targets presented during the AB (Popple and Levi (2007); Vul, Nieuwenstein,

and Kanwisher (2008); see also Section 7.1.2). In line with this argument, our results show that the EEG correlates of attentional selection and working memory encoding are more 'jittered' in time for targets presented during the AB when compared to targets presented outside the AB. The notion of increased temporal variance in target processing during the AB is inherent to the theoretical framework underlying the $ST^2$ model. The $ST^2$ model suggests that working memory encoding is serial during the AB, hence processing of the second target is delayed until the consolidation of the first target has completed. Consequently, the time it takes to process the first target determines the time point at which the second target can be processed.

In Chapters 6 and 7, this thesis provides a theoretical contribution to assessing current theories of temporal attention against experimental data. Whereas the EEG results presented in Chapter 6 are in contrast with competition-based accounts of the AB (such as the theories described in Section 3.1.3), the EEG results from both Chapters 6 and 7 provide evidence for a notion of serial working memory encoding during the AB, which supports the theory underlying the $ST^2$ model.

**How the attentional blink modulates the influence of target strength on conscious perception**

In Chapter 8, we investigate how bottom-up target strength and the availability of attention influence whether a target can be consciously perceived in RSVP. To this aim, we compare the P3 component, which is assumed to be the EEG correlate of conscious perception for targets in RSVP, for targets presented individually (i.e. outside the AB) to targets presented during the AB. We find that target perception outside the AB is mainly dependant on target strength and that target strength influences the profile of the P3 component. For targets inside the AB, however, conscious perception is determined by the availability of attention and the P3 component is not affected by target strength. We show how the $ST^2$ model cannot account for these results and propose a modified theory that can explain both our experimental results and also previous findings by Dehaene and colleagues (Sergent & Dehaene, 2004; Sergent et al., 2005; Del Cul, Baillet, & Dehaene, 2007).

### 1.2.2 Applying cognitive electrophysiology to neural modelling

As the methodological contribution of this thesis, we propose a direct connection between computational modelling of cognition and cognitive electrophysiology. We further develop a computational model of working memory and temporal attention (Bowman & Wyble, 2007) to provide us with predictions about EEG data. The computational model employs a neural network architecture, which, although used in an abstract and high-level manner, is based on neurophysiologically plausible processes. Each unit in the network exhibits certain neural activation dynamics. The hypothesis to be tested in this thesis is whether summations of these activation traces - which we refer to as *virtual ERPs* - can provide meaningful predictions and explanations of the EEG traces recorded from human participants. In the following, we describe how the virtual ERP technique as the methodological contribution of this thesis will be assessed with regards to a number of criteria.

#### An emergent property

The computational model that is used to generate virtual ERPs was originally designed to simulate behavioural data. We propose to take this model and - despite making only minimal changes to its parameters compared to the originally published version - use the model to explore an additional dimension of experimentation, namely EEG. Virtual ERPs can thus be classified as an *emergent property* (Goldstein, 1999) of the computational model. Although the model was designed to do one thing (i.e. replication of behavioural results), we propose to use this same model to make predictions about the human EEG.

#### Assessing cognitive models

We propose that virtual ERPs provide a powerful tool for assessing the computational model. The virtual ERP technique utilises the neural architecture of the model and allows us to directly analyse the activation dynamics of the neural network.

When using computational models of cognition to replicate behavioural data, the main emphasis is on the output of the model (e.g. Cohen, Romero, Servan-Schreiber, & Farah, 1994). In other words, the processes underlying the model might be neglected, as long as

the values produced by the simulation are a good replication of the human data. Consequently, it is sometimes difficult to determine the advantage of neural models over other forms of computational modelling of cognition, for instance using closed form equation models (O'Reilly & Munakata, 2000; Levine, 2000). Specifically, if both types of models are equally able to replicate the behavioural data, a model employing closed form equations might be preferred to a neural model because it can be more straightforward and thus less complex.

The virtual ERP technique that we propose provides insights into the neural activation dynamics during the process of a simulation and emphasises the advantage of using neural networks to model cognition. The virtual ERP can shed light on the 'blackbox' of intermediate processing that occurs while the results of a simulation are being generated. Consequently, we can not just analyse the output from the model but the virtual ERPs also illustrate the processes that lead to a given set of simulation results.

#### An additional method of mining EEG

Virtual ERPs can complement recent developments in the 'information-based approach to modelling electroencephalographic (EEG) dynamics' (Makeig, Debener, et al., 2004). We propose to use the virtual ERP method to find theoretical explanations for the effects that certain experimental manipulations have on the human EEG profile. Accordingly, if we find that a manipulation has the same effect in both human and virtual ERP, we can dissect the virtual ERP by investigating the activation dynamics of the model and find out what produced the effect in the model. These findings can then be used to make conclusions about what caused the experimental effects in the human data.

### 1.2.3 Using temporal attention research for computer systems design

In Chapter 9, we discuss a practical application for findings from research into temporal attention and EEG. We show how the design and behaviour of computer systems might be improved by taking into account, first, the nature of the human attentional system and, second, current brain states expressed by the EEG signal of the user interacting with the computer system.

## 1.3 Publications

The experimental findings and discussions presented in this thesis have partly been published elsewhere. The concept of virtual ERPs evolved from the initial description published in Bowman, Wyble, Chennu, and Craston (2008) over a number of presentations at conferences (Craston, Wyble, & Bowman, 2006, 2007) to a journal publication (Craston, Wyble, Chennu, & Bowman, 2009). The experimental data presented in Chapters 5 and 6 extend the analysis presented in Craston et al. (2006). Craston et al. (2006) also contained a first attempt at the single-trial ERPimage analysis of the P3 component presented in Chapter 8 and a preliminary analysis of comparing target processing in standard RSVP to skeletal paradigms (Chapter 5). Extending the results presented in Craston et al. (2007), the findings on target consolidation during the attentional blink from Chapter 6 have been published as an article in the Journal of Cognitive Neuroscience (Craston et al., 2009). A first analysis of attentional processes during the attentional blink was presented in Chennu, Craston, Wyble, and Bowman (2008) and the further analyses presented in Chapter 7 are in preparation for journal publication (Chennu, Craston, Wyble, & Bowman, in revision). The discussion of practical implications of EEG research for the design of computer systems discussed in Section 9 is based on the work published as a technical report (Wyble, Craston, & Bowman, 2006).

### 1.3.1 Personal contribution to the collaborative research

Some of the work presented in this thesis originated from collaborative research. The $ST^2$ model, as published in Bowman and Wyble (2007), was designed and implemented by Brad Wyble and Howard Bowman. I took the existing neural network implementation of the model and added the functionality to extract and visualise neural activation traces, which was required to generate virtual ERPs. I devised the methodology that is used to generate the virtual ERPs (as described in Section 4.3). The virtual ERPs allow a detailed investigation of the neural dynamics of the $ST^2$ model and I was subsequently able to improve the architecture of the neural implementation of the model in a number of ways (see Section 4.4 for details).

The two EEG experiments, of which the results are presented in Chapters 5, 6, 7, 8

and 9, were designed by me and conducted in collaboration with Srivas Chennu. Whereas the EEG data presented in Chapters 5, 6 and 8 were analysed and interpreted by me, the analysis for Chapter 7 was performed in collaboration with Srivas Chennu. The analysis presented in Chapter 9 was performed in collaboration with Brad Wyble.

## 1.4 Organisation

### 1.4.1 Part I

Following this general introduction, Part I continues with a review of the literature relevant to the work presented in this thesis. First, we provide a basic introduction into cognitive electrophysiology using EEG. This is followed by a brief overview of computational modelling of cognition. Finally, we review the literature on visual processing under high temporal demands. Particular emphasis is on the attentional blink phenomenon.

### 1.4.2 Part II

Part II is concerned with introducing a novel methodological technique of establishing a direct connection between neural modelling of cognition and cognitive electrophysiology. We commence with a review of current informal and formal theories of the attentional blink. With respect to formal theories of the attentional blink, we assess their ability to simulate electrophysiological data. As it is the basis for the modelling work presented in this thesis, we continue by providing a detailed description of the Simultaneous Type Serial Token ($ST^2$) model (Bowman & Wyble, 2007). Following this, we describe in detail how virtual ERPs are generated from the $ST^2$ model.

### 1.4.3 Part III

Part III presents the experimental results of this thesis. The data are analysed and interpreted with respect to theories of the attentional blink. This part also evaluates the virtual ERP technique as a contribution to scientific research.

### 1.4.4 Part IV

The thesis concludes with a summary and an assessment of the extent to which this project has provided the scientific contributions proposed in the introduction. Furthermore, we describe a potential practical application for insights from temporal attentional research when designing computer systems and discuss potential future directions of research inspired by the work presented in this thesis. The appendix contains a description of the methods employed in the experiments that were conducted for this thesis.

# Chapter 2

# Literature review

The following chapter reviews the literature to provide a theoretical background for the work presented in this thesis. We begin with a general description of EEG research and then describe the EEG analysis techniques that are employed in this thesis. We continue with a brief description of computational modelling of cognition and, finally, provide an overview of research on the attentional blink.

## 2.1 Electroencephalography

The neurophysiological measurement of electrical brain activity is known as electroencephalography. Richard Canton, an English physician, was the first to discover electrical currents being elicited from rabbit and monkey brains in 1875 (see Swartz, 1988). These electrodes were inserted directly into the animal brain and suggested a correlation between the electrical signal and neural activity in the brains of these living organisms. In 1924, the German neurologist Hans Berger discovered one could also record electrical currents from electrodes placed on the human scalp (Berger, 1929). This finding provided a means for recording brain activity by non-invasive measure, thus opening up the possibility of studying brain activity in vivo in humans.

The electrical activity recorded from electrodes placed on the scalp is known as the electroencephalogram (EEG). As the electric potentials produced by neurons in the brain are relatively weak, a powerful EEG amplifier is needed to make the signal visible. In addition to brain-related activity, the EEG amplifier will also record electrical signals from

the environment, such as line power noise and muscular activity from the participant's body. Hence, EEG research requires sophisticated analysis techniques that increase the signal-to-noise ratio in order to extract a signal from irrelevant background activity.

The experimental work presented in this thesis (in Chapters 5, 6, 7 and 8) contains behavioural results as well as EEG data. In order to interpret the EEG data, we employ a number of analysis techniques and these are described in the following sections.

### 2.1.1 Event-related potentials

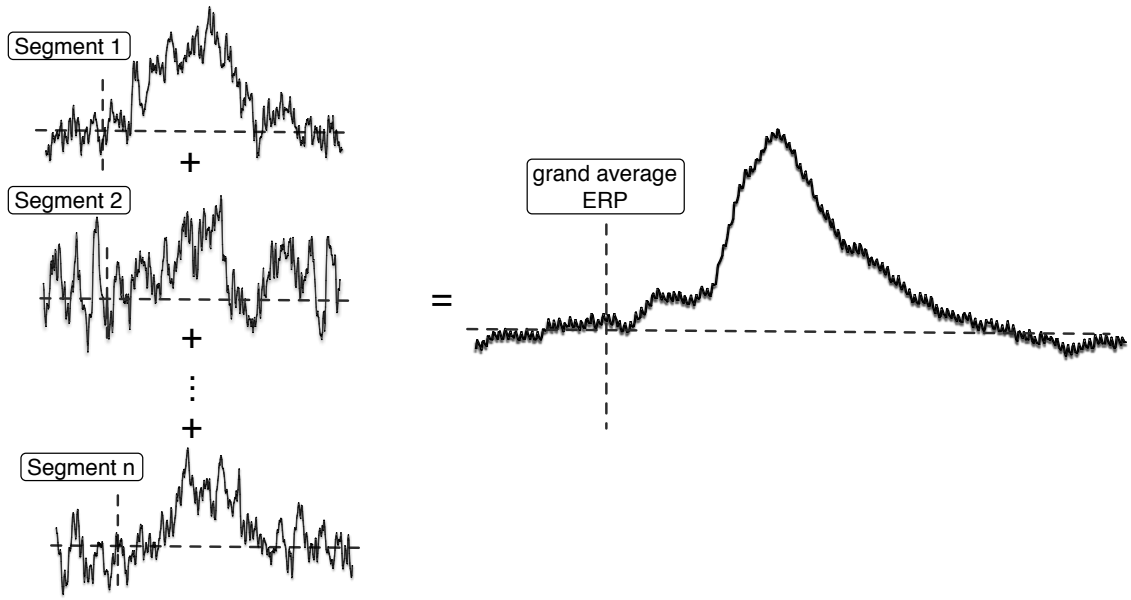


*Figure 1* Averaging segments of raw EEG to extract event-related potentials. Positive is plotted upwards.

As seen in Figure 1, event-related potentials (or ERPs) are generated by averaging across multiple segments of EEG activity time locked to an externally generated event. The averaging process increases the observable signal by removing ongoing non-time locked EEG activity, which is treated as background noise. The resulting ERP waveform contains a number of positive and negative deflections, which are referred to as ERP components. A number of ERP components have been associated with cognitive processes during various experimental tasks. Through correlational evidence researchers can infer the consequences of experimental manipulations on the participant's brain activity, as reflected by the EEG.

### ERP components

EEG is employed as a research tool in a wide range of scientific areas, which has led to a large number of ERP components being reported. Since an exhaustive review is beyond the scope of this work, we focus on ERP components that are relevant for the study of the human attentional system with respect to vision and are used in the analyses of EEG data presented in this thesis.

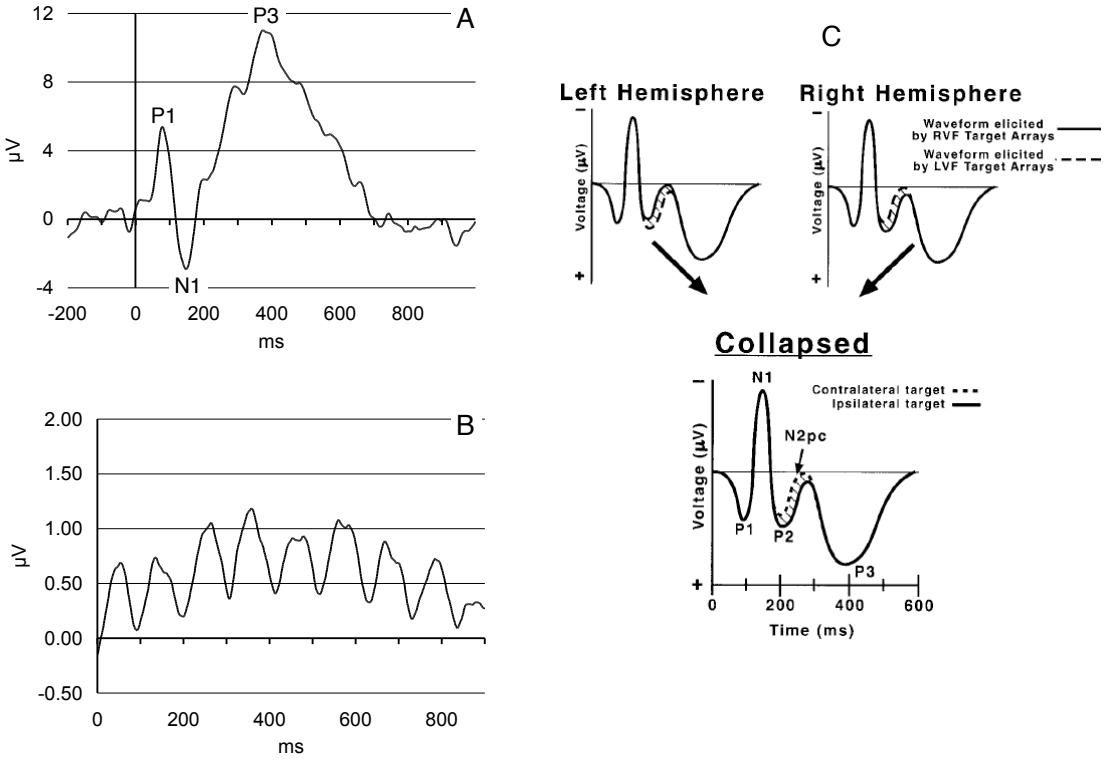


*Figure 2* Panel A: A sample ERP showing the P1, N1 and P3 ERP components. Positive is plotted upwards. Panel B: A sample ssVEP wave oscillating at 10Hz. Positive is plotted upwards. Panel C: An illustration of how to extract a lateralised ERP component, such as the N2pc. Figure adapted from Woodman and Luck (2003). In this figure negative is plotted upwards.

**Early sensory processing** The P1 and N1 ERP components (Figure 2A) - commonly associated with early sensory processing (Hillyard, Vogel, & Luck, 1998) - are observed for individually presented items. The components' notations indicate that they are often the first positive and negative deflections of the ERP. P1 and N1 typically occur around 100-200ms after stimulus presentation. Repeatedly presented items, on the other hand, evoke

the steady-state Visually Evoked Potential (ssVEP) (Figure 2B), a wave oscillating at the same frequency as the presentation rate of the items displayed to the observer (Müller & Hubner, 2002; Müller et al., 1998; Di Russo, Teder-Sälejärvi, & Hillyard, 2003).

**Attentional selection** The N2pc ERP component has been described as a correlate of attentional selection when subjects are required to detect targets among irrelevant distractor items (Luck & Hillyard, 1994b; Eimer, 1996; Hopf et al., 2000). The N2pc occurs around 150-300ms post-stimulus presentation and is a lateralised negative deflection of the ERP. Hence, the N2pc is only visible in the difference waveform between EEG activity from ipsilateral and contralateral electrodes (see Figure 2C).

**Working memory consolidation** The P3 (or P300) is the third positive peak of the ERP and is often the most distinctive ERP component, occurring between 300 and 600ms post-stimulus presentation depending on the type of task (Figure 2A). The P3 is evoked most strongly by a rare event amongst a sequence of frequent items, through so-called oddball tasks. The exact cognitive processes underlying the P3 have been subject to much debate (for an extensive discussion see Donchin and Coles (1988) and Verleger (1988)). With respect to the work presented in this thesis, however, the P3 component is seen as a correlate of consolidating items into working memory (Donchin, 1981; Vogel, Luck, & Shapiro, 1998). This is supported by the finding that correctly reported targets evoke a P3, whereas a P3 component is not observed for targets that cannot be reported (Kranczioch, Debener, & Engel, 2003).

**Semantic processing** The N400 component is commonly associated with semantic processing (Kutas & Hillyard, 1980; Kutas & Van Petten, 1994; Van Berkum, Hagoort, & Brown, 1999) and describes a negative deflection of the ERP occurring at approx. 400ms post-stimulus presentation. The N400 is evoked by semantic incongruency, for instance by unexpected words at the end of a sentence. A common example is 'He spread the warm bread with socks', which will evoke a large N400 component time locked to the word 'socks'. If, on the other hand, the expected sentence ('He spread the warm bread with butter') is presented to the participant, the negative deflection in the N400 range is reduced.

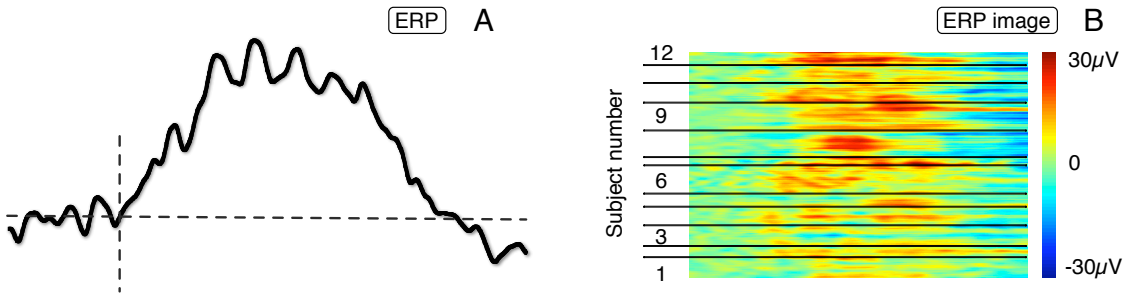


*Figure 3* Panel A: One-dimensional ERP plot containing EEG activity averaged over time. Positive is plotted upwards. Panel B: ERPimage plot: Time is plotted on the x-axis and individual EEG epochs are plotted on the y-axis. Voltage values are displayed using a colour coded scale and the data is smoothed along the y-axis to enhance the visual signal-to-noise ratio.

### ERPimages

ERP analysis is a powerful tool for experimental research in cognitive psychology. However, when generating an ERP average (see Figure 3A), the continuous raw EEG is reduced to a one dimensional dataset displaying a sequence of voltage fluctuations over time. The averaging process extracts EEG activity that is time locked to the stimulus, whereas the rest of the signal is treated as irrelevant background noise. The problem with this approach (which is common to the averaging process in general) is that although one can extract the overall trend of the data, information that is specific to individual observations - but not present throughout the data - is lost. A typical example is a data set consisting of very high and very low values. The average value would be half way between the high and the low value and would not accurately reflect the variance of the underlying data.

Figure 3B shows a so-called ERPimage plot (Makeig, Debener, et al., 2004; Delorme & Makeig, 2004). Whereas the grand average ERP in Figure 3A averages across individual epochs of raw EEG activity, the ERPimage displays the raw EEG epochs stacked on top of each other. Voltage values per time point (time is plotted on the x-axis) are expressed by means of a colour scale where blue indicates negative and red represents positive values. The y-axis contains the individual epochs of raw EEG activation. The epochs in an ERPimage can be sorted by a number of criteria, for instance subjects (as seen in Figure 3B) or phase of a given frequency (see Makeig, Delorme, et al. (2004)). In order to enhance the visual signal-to-noise ratio, the ERPimage can be smoothed across epochs using a moving average

window. The ERPimage can thus be used to investigate differences between individual trials, which are not visible in the grand average ERP waveform (see also Chapter 7 and 8).

### 2.1.2 ERSP and ITC time-frequency analysis

Each channel of the EEG is a continuous waveform consisting of waves oscillating at various frequencies. Average power per frequency can be analysed by transforming the ERP from the time to the frequency domain, for instance by means of a Fast Fourier Transform (FFT) or wavelet decompositions.

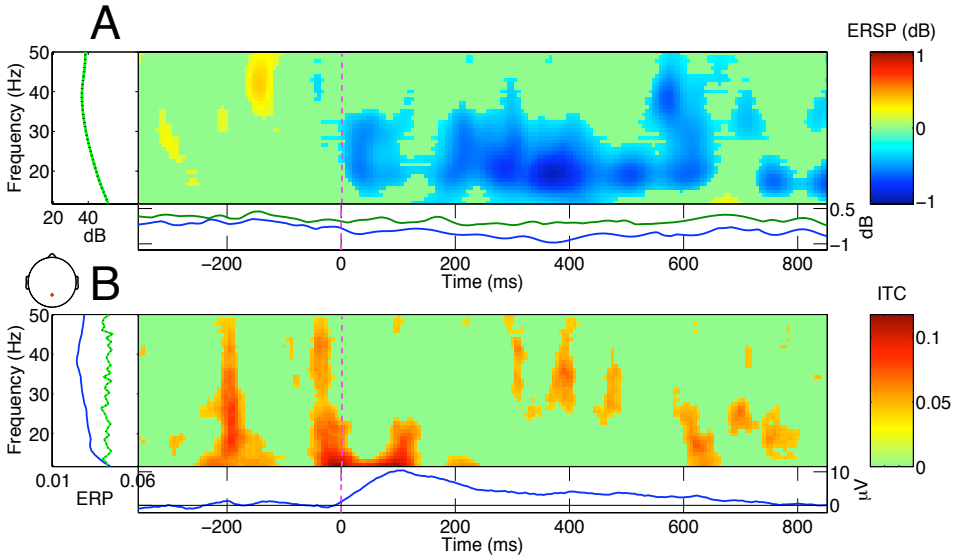


*Figure 4* Panel A: Event-related spectral perturbation (ERSP) plot. Time is plotted on the x-axis and the y-axis contains the frequency bands of the specified range. Power is displayed using the colour scales specified to the right of the plot. The waveform to the left shows the baseline average power per frequency band. The waveform below depicts the highest and lowest mean power values across all frequencies relative to baseline at each time point. Panel B: Inter-trial coherence (ITC) plot. Time is plotted on the x-axis and the y-axis contains the frequency bands of the specified range. ITC phase coherence values are displayed using the colour scales specified to the right. The plot to the left depicts the minimum and maximum mean ITC values per frequency across all time points. The waveform below shows the grand average ERP waveform.

Once the EEG data has been transformed to the frequency domain, it can be analysed using various techniques. One of these is event-related spectral perturbation (ERSP; Makeig, Debener, et al., 2004; see Figure 4A for an example plot), which displays the

amount of power in a given set of frequency bands over time after averaging across the raw EEG epochs time locked to the presentation of a stimulus. Time is represented on the x-axis, the y-axis contains the frequencies in the specified range and the amount of power is displayed using a colour scale. Consequently, increased 'redness' at a particular datapoint indicates a larger amount of power in that specific frequency band at a given point in time.

Another time-frequency analysis technique is inter-trial phase coherence analysis (ITC; see Figure 4B for an example plot), which measures the extent to which the phase of EEG activity time locked to the presentation of a stimulus is correlated across a set of EEG epochs (Delorme & Makeig, 2004; Makeig, Debener, et al., 2004; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). Since the ITC analysis is performed on a specified range of the raw EEG's frequency spectrum, it provides a measure of the phase synchronisation in different frequency bands. ITC is typically calculated as a dimensionless number normalised to a value between 0 and 1. An ITC value of 0 indicates a complete lack of phase coherence across the set of epochs being analysed, whereas an ITC value of 1 signifies that the phase changes are perfectly time locked to the stimulus. ITC can then be visualised in a time-frequency colourmap with time, with respect to the stimulus onset, along one dimension, and frequency along another. Each datapoint in the plot is coloured according to the amount of ITC observed at a particular time and frequency across all epochs in an experimental condition. As indicated by the colour scale to the right, increased 'redness' depicts a larger ITC value, and thus more phase synchronisation across epochs.

The various types of EEG oscillations are commonly referred to by their frequency bands, namely Delta for frequencies of 4Hz and below, Theta for 4-8Hz, Alpha for 8-14Hz, Beta for 14-30Hz, Gamma for 30Hz and above. Similarly to ERP components, EEG oscillations in certain frequency bands have been associated with cognitive processes. Alpha is the most prominent oscillation in the EEG signal and is most pronounced at posterior electrode locations. Alpha waves are often strong enough to be observed in the raw EEG, for instance if participants close their eyes or are in a relaxed or inattentive state (Berger, 1929). Oscillations in higher frequency bands, however, require a time-frequency transformation to be made visible. Within the higher frequency bands, beta and gamma oscillations have been connected to attentional selection and the conscious identification of stimuli (Gross et al., 2004; Engel & Singer, 2001; Fell, Fernandez, Klaver, Elger, & Fries, 2003; Kranczioch,

Debener, Herrmann, & Engel, 2006).

### 2.1.3 EEG source localisation

Despite extensive research for almost a decade, the neural substrates of EEG are still to be fully explained. What is known, however, is that the difference in electric charge between the dendrite and the postsynaptic cell body of an active neuron creates an electric dipole. To generate a signal that is strong enough to be registered by the EEG amplifier, a population of neurons have to be active together and also spatially aligned, which causes the individual dipoles to summate. Cortical pyramidal neurons have long-range connections and are aligned perpendicular to the cortex, which is why these neurons are assumed to be a major contributor of the human EEG (Luck, 2005).

The mapping of EEG activity measured at the surface of the scalp to individual brain regions is constrained by two mathematical dilemmas, namely the forward problem and the inverse problem. The forward problem arises when calculating a signal (the potential at the scalp) from a given number of sources (electrical current generated by a neuron). The extent of the forward problem is determined by the amount of resistance and other signal distortion between the source and the location where the signal is measured. Thus, in order to approximate a solution to the forward problem in EEG, one needs to calculate the conductance of body tissue between the neuron and the electrode. The volume conductor, which in the case of EEG is the human head, has been modelled using various methods (Kavanagk, Darcey, Lehmann, & Fender, 1978; Oostendorp & van Oosterom, 1989) and, hence, a number of estimated solutions to the forward problem exist.

The second mathematical dilemma in EEG source localisation is expressed by the inverse problem, i.e. localising the underlying sources from a given signal (von Helmholtz, 1853). The inverse problem is exacerbated if the number of underlying sources is unknown, as is the case with EEG. However, one can approximate the location of an equivalent dipole instead of localising individual neurons. The equivalent dipole is assumed to reflect the origin of the pattern of neural activity recorded by the EEG.

The dipole fitting approach (Oostenfeld & Delorme, 2007) estimates a solution to the inverse problem by performing a nonlinear search. When using this technique, the algorithm is provided with an initial guess for a dipole location. The algorithm then calculates the

error of the initial location using the volume conductor model (which - as described in the previous paragraph - provides an estimated solution for the forward problem) and makes slight changes to the parameters until the error has been minimised.

## 2.2 Computational modelling of cognition

How does cognition and consciousness emerge from the brain? This question, often referred to as the *Mind-Body Problem* (see Taylor (2008) for an overview), has been the subject of much debate since the times of Plato, Aristotle and other ancient Greek and Eastern philosophers. For a long time, this debate was dominated by dualists, which propose that the mind (and hence cognition) and the body (i.e. the brain) are distinct entities (e.g. Descartes, 1641). However, such an account is unsatisfactory from a scientific point of view as it seems to raise more questions that it answers (for instance, in which way do the mind and body interact?). The alternative is a reductive materialist approach, which avoids any supernatural explanations. Instead, it is assumed that there is only the body and that cognition must therefore emerge from biological mechanisms occurring in the brain.

Since the late 19th century, psychologists made progress in studying the mind by using human behaviour to understand cognition. Parallel to this, neuroscientists and neurologists established a detailed characterisation of brain anatomy. Nevertheless, up until very recently, these two areas of research have remained separate and there has not been much interaction between them. The relatively new field of cognitive neuroscience bridges this gap in order to find a neurological correlate of human cognition (Gazzaniga, 2004). In addition to measuring behavioural performance, cognitive neuroscience employs brain-imaging techniques to visualise participants' brain activity while performing a given task. Through these experiments, researchers have been able to associate brain areas with cognitive functions.

However, the results gained from these experiments need to be integrated into theoretical frameworks in order to be accurately interpreted. This is often done by devising models of cognitive processes to describe a theory. Although a model is always a simplification of the real world, if designed in the right way, it will leave less room for ambiguity than an abstract theory (Frigg & Hartmann, 2006). Furthermore, models can be used to make

predictions about related experimental results and provide a means for validating the underlying theory (Popper, 1959). Consequently, cognitive modelling is an important part of cognitive neuroscience.

### 2.2.1 Symbolic models

In traditional psychology, cognitive theories are often described using informal models, which employ iconic representations such as boxes, arrows, characters or logical operators. Such models provide a means for breaking down the cognitive theory into various subsystems and describing how the individual components of the system interact under various circumstances (see Barnard (1999) for an example of such a model). These informal models describe cognition in terms of functional mental processes and structural subsystems, while often abstracting away from the neurological basis underlying cognition.

Another approach uses mathematical equations to model cognition (e.g. Kalidindi & Bowman, 2007). These equation models often refrain from describing the structural properties of a cognitive system, but they can be implemented using computational methods. Hence, whereas box-and-arrow models illustrate the structural properties of a cognitive system but commonly only generate rather informal predictions, equation models provide formal predictions at the expense of a structural description. Equation models describe the cognitive theory in a non-ambiguous manner and the predictions derived from equation models can be used to validate the cognitive model against experimental results. Both box-and-arrow and equation models use symbolic descriptions, i.e. language, numbers or schematic signs, to describe a cognitive theory. It is clear, however, that the brain does not directly work in this way. Rather, cognition seems to emerge from the brain by means of the exchange of electric signals between neurons in the brain.

### 2.2.2 Subsymbolic models

Neural network models of cognition address this issue by taking inspiration from the biological architecture of the brain and model cognition by simulating the mechanisms that are assumed to occur in the brain (see e.g. Garson (2007); O'Reilly and Munakata (2000)). In analogy with neurons in the brain, neural networks contain layers of relatively simple

nodes (see Figure 5 for an example of a very simple neural network). Weighted connections between the nodes (illustrated by the arrows between the nodes in Figure 5) correspond to synaptic projections in the brain and transfer activation between the nodes. As they do not contain symbolic representations, neural network models are commonly referred to as subsymbolic models (Fodor & Pylyshyn, 1988).

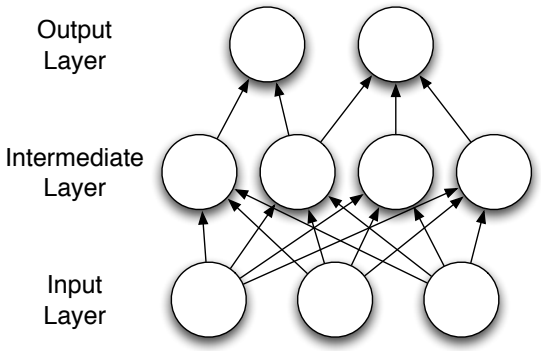


*Figure 5* Example of a simple neural network containing three layers. In such a neural network, the input layer will typically receive an input signal in the form of some activation pattern. This activation feeds through the hidden layer where it is modulated according to the weight values of the connections between the layers. The output layer then provides the user with an activation pattern that can be interpreted.

In a typical neural network (as shown in Figure 5), an input layer receives activation from some external source. The activation then feeds through weighted connections and causes nodes in the following layers to accumulate activation. The example in Figure 5 contains only one intermediate layer, however, a neural network will typically contain multiple intermediate layers. Once the nodes in a layer reach a given threshold, they 'fire' and activation gets passed on to nodes in subsequent layers via the corresponding connections. Thus, similar to neurons in the brain, neural network nodes act as detector units (O'Reilly & Munakata, 2000) and respond once a signal crosses a given threshold. The input signal is transformed while the activation feeds through the intermediate layers of the network until it reaches an output layer where the signal is interpreted. The neural network thus functions as an information processing system.

The $ST^2$ model, which is described in detail in Section 3.3, is such a neural network model. As we will show in Section 3.3, the $ST^2$ model provides a theoretical description of

temporal attention and working memory in humans. The $ST^2$ model also has a formal computational implementation and can be used to run simulations, which generate predictions that can be validated against experimental data.

## 2.3 Visual processing under high temporal demands

In daily life, humans are constantly exposed to continuous streams of sensory input. The attentional system acts as a filtering mechanism that distinguishes relevant from irrelevant stimuli. Most of the time, the visual system performs remarkably well at configuring a percept of the world that allows us to navigate and manoeuvre in a safe manner. Under certain circumstances, however, observers consistently fail to detect stimuli or can be fooled into misperception, such as optical illusions.

The failure to detect stimuli can have severe consequences. When navigating in car traffic, for instance, humans are required to pay constant attention to their environment. Also, modern safety-critical technology often requires humans to react to messages and signals that are presented on computer screens for brief periods of time only. Therefore, the design of efficient computer systems requires detailed knowledge about the nature of human attention (see Chapter 9 for a further discussion of this issue).

This thesis is concerned with investigating the nature of temporal attention in humans. In order to establish the background for the work that is presented in Chapters 5 to 8, the following sections review the relevant literature on temporal attention and provide a detailed description of previous research investigating the attentional blink.

### 2.3.1 Rapid serial visual presentation

The rapid serial visual presentation (RSVP) paradigm has been extensively used to explore the temporal properties of the human visual system (Lawrence, 1971; Broadbent & Broadbent, 1987). Stimuli in an RSVP stream are often alphanumeric characters, such as letters and digits (Weichselgartner & Sperling, 1987; Reeves & Sperling, 1986), presented at rates of 10-20 items per second in the same spatial location. At this speed, observers perceive only fleeting mental representations of individual items, as each item masks its predecessor.

Nevertheless, the visual system does a remarkably good job at detecting target items. Various studies have shown that participants are able to extract complex semantic information from pictures (Potter, 1976) or words (Barnard, Scott, Taylor, May, & Knightley, 2004) despite the fast presentation rate of RSVP. Figure 6 shows two different types of RSVP experiments, where observers are asked to report target letters that are embedded in a stream of rapidly presented distractor items.

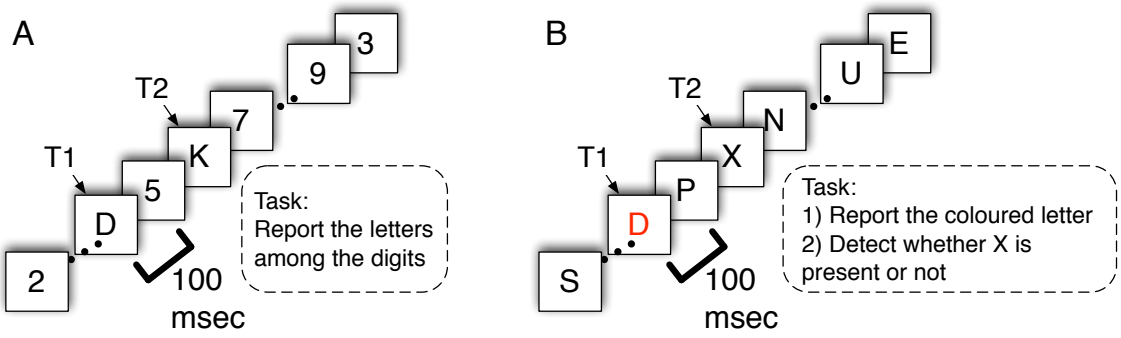


*Figure 6* The RSVP paradigm that is typically used in attentional blink experiments. Panel A: A letters-in-digits attentional blink task as used in Chun and Potter (1995). Panel B: A colour-marked attentional blink task as used in Raymond et al. (1992)

### 2.3.2 The attentional blink

Under certain circumstances, however, observers consistently fail to detect target items when presented in RSVP paradigms such as the ones depicted in Figure 6. The attentional blink (AB; Broadbent & Broadbent, 1987; Raymond et al., 1992; Chun & Potter, 1995) is a particularly striking example of this limitation. The AB describes a finding that detection of a second target (T2) is severely impaired if T2 follows an identified first target (T1) within approx. 600ms (the T2|T1 curve in Figure 7). This alleged blink of the 'mind's eye' (Raymond et al., 1992) was initially thought to reflect a fundamental limitation of visual perception in humans.

#### Early or late bottleneck?

Subsequent research, however, suggests that the AB is by no means absolute. As seen in Figure 7, even during the deepest part of the AB, a number of T2s are detected, i.e.

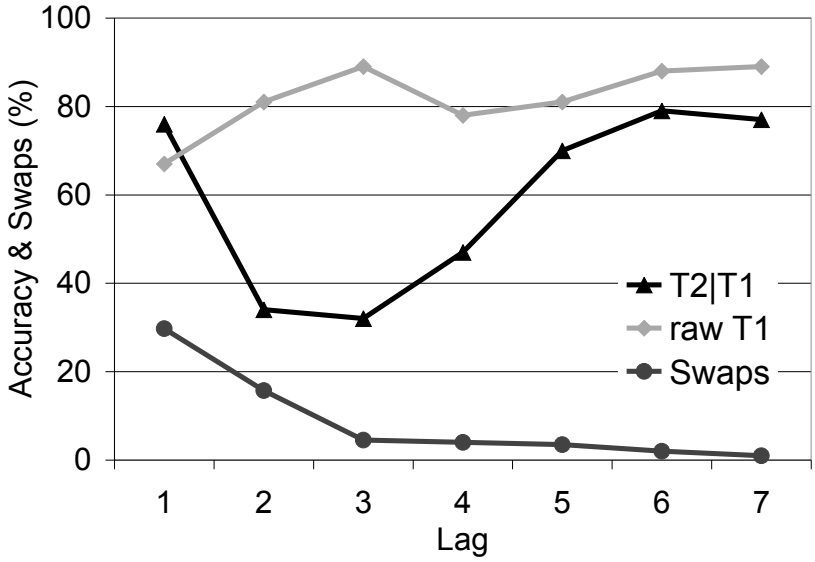


*Figure 7* The attentional blink. T2 accuracy is conditional on correct report of T1 (T2|T1), raw T1 accuracy and percentage of swaps. The data is from Chun and Potter (1995).

performance is never at zero (Chun & Potter, 1995). Furthermore, experimental evidence suggests that even if T2 cannot be consciously reported during the AB, it is nevertheless processed to a semantic level. Accordingly, ERP components associated with sensory (Vogel et al., 1998; Sergent et al., 2005) and also semantic (Luck, Vogel, & Shapiro, 1996; Rolke, Heil, Streb, & Hennighausen, 2001) processing have been found to remain present for T2s that are missed during the AB. In addition, it has been found that T2s that are missed during the AB can still prime a subsequent item (Shapiro, Driver, Ward, & Sorensen, 1997). The P3 component, associated with working memory encoding, is present for T2s that are reported during the AB but much reduced for missed T2s during the AB (Kranczioch et al., 2003; Sergent et al., 2005). These findings suggest a late stage bottleneck, i.e. that targets presented during the AB are initially processed but then often lost before they can be consolidated into working memory.

**'Sparing' at lag 1**

Many studies exploring the AB employ RSVP paradigms where one target is of a different category than distractors (e.g. a vowel among consonant distractors) and the other target

is colour marked (see Figure 6B). The task switch between category search for one target and colour detection for the other, however, has been suggested to be a potential confound (Potter, Chun, Banks, & Muckenhoupt, 1998). A letters-in-digits task (Figure 6A), where all items have the same colour and both targets are distinguished from distractors by category, has been claimed to be a 'purer' version of the AB paradigm (Chun & Potter, 1995).

As seen in Figure 7, observers do remarkably well at detecting a T2 that is presented in immediate succession to T1. Such 'lag 1 sparing' (Chun & Potter, 1995) is observed if targets appear in the same spatial position (Visser, Zuvic, Bishof, & Di Lollo, 1999) and is particularly strong if there is no task switch between T1 and T2 (Chun & Potter, 1995).

An RSVP stream at twice the typical presentation rate (20 items per second) has been shown to elicit 'lag 2 sparing', even though there is a distractor in the T1+1 position (Bowman & Wyble, 2007). Thus, lag 1 sparing seems to be about time rather than sequential position. Furthermore, a distractor presented in the lag 1 position is more likely to prime a following target compared to other distractors (Chua, Goh, & Hon, 2001). Hence, the processing of items (both targets and distractors) that follow T1 within approximately 150ms seems to be enhanced.

Recently, it has been suggested that lag 1 sparing comes at the cost of there being a trade-off effect if both targets appear in close temporal proximity (Potter, Staub, & O'Connor, 2002). As shown in Figure 7, good T2 performance comes at the cost of decreased T1 performance at lag 1 (Chun & Potter, 1995). Figure 7 also shows that there is a higher percentage of swaps at lag 1, in that participants report the identity of both targets correctly but often confuse the order that the targets were presented in (Chun & Potter, 1995; Hommel & Akyürek, 2005; Bowman & Wyble, 2007). The loss of temporal distinctiveness is further increased for more complex targets. If targets are letter pairs, for instance, in addition to confusing the order between complete targets at lag 1, subcomponents of targets (i.e. letters) migrate between T1 and T2 (Bowman & Wyble, 2007).

### Circumventing the attentional blink

A number of recent studies have shown that the AB can be eliminated during periods where it was thought to be strongest, i.e. between lags 2-5. These findings present a

particular challenge for theories of the AB claiming that the AB reflects a fundamental capacity limitation of the human attentional system. The finding is referred to as *spreading*

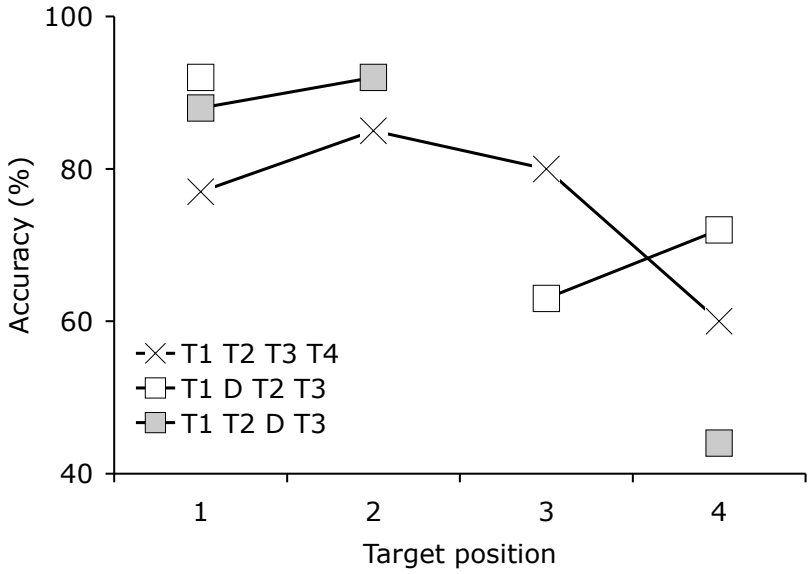


*Figure 8* Accuracy per target position for (a) T1 T2 T3 T4: No AB if four targets are presented in a row. (b) T1 D T2 T3: If T1 is followed by a distractor, then T2 is 'blinked', however, T3 is 'spared'. (c) T1 T2 D T3: T2 is presented following T1 and is 'spared', T2 and T3 are separated by a distractor and T3 is 'blinked'. The data from curve (a) is from Nieuwenstein and Potter (2006). The data in curves (b) and (c) are from Olivers et al. (2007).

*the sparing* (Olivers et al., 2007; see also Di Lollo, Kawahara, Ghorashi, and Enns (2005); Kawahara, Kumada, and Di Lollo (2006) for a similar account) where participants are able to report targets despite being presented during the deepest part of the AB. These paradigms employ RSVP streams with multiple targets (T1, T2, T3 and T4). If all targets are presented in a row (curve (a) in Figure 8), accuracy at detecting the targets is reasonably high, i.e. no AB occurs. If the targets are intermixed with distractors, each target seems to open up a new 'sparing window', which benefits detection of an immediately following target, as long as those two targets are not separated by distractors (i.e. T3 in curve (b) and T2 in curve (c) of Figure 8). However, if a target is separated by distractors from the previous target, an AB occurs and performance at detecting that target suffers (i.e. T2 in curve (b) and T3 in curve (c) of Figure 8).

Furthermore, it seems as if the mental state of the observer is a critical factor for determining whether an AB occurs or not. If participants are distracted by a concurrent irrelevant task (Olivers & Nieuwenhuis, 2006) or visual background noise during the experiment (Arend, Johnston, & Shapiro, 2006), the AB is reduced. The AB was also found to be reduced if participants listen to music while viewing the RSVP stream (Olivers & Nieuwenhuis, 2005; Ho, Mason, & Spence, 2007).

In most AB experiments, there is variance in the strength of the AB between subjects and, in some experiments, a considerable number of participants show no AB at all. Whereas it seems that through training (Ruthruff, Johnston, Van Selst, Whitsell, & Remington, 2003) or meditation (Slagter et al., 2007) one can learn how to allocate attentional resources more efficiently, some 'non-blinkers' (Martens, Munneke, Smid, & Johnson, 2006) are seemingly able to adopt this beneficial strategy without prior experience at the task.

**The attentional blink and masking**

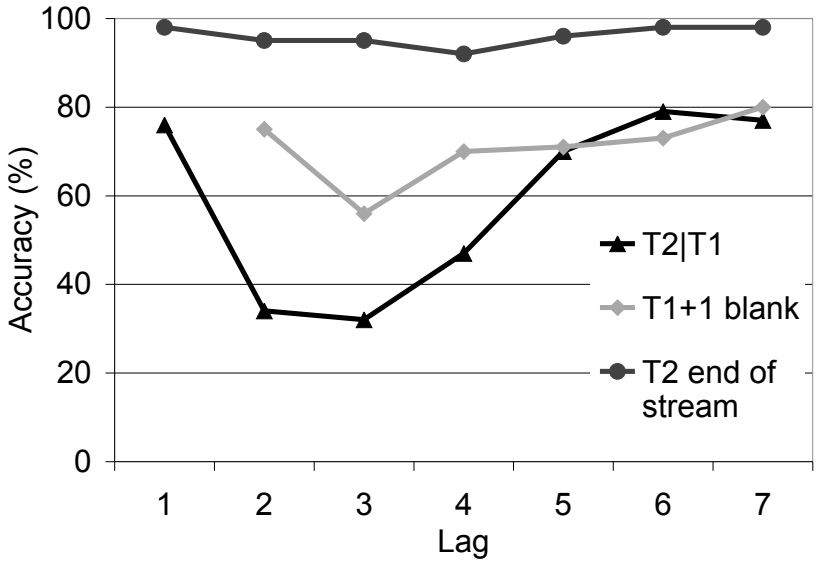


*Figure 9* The AB is reduced if the distractors following each of the targets (T1 and T2) are removed. T2 accuracy conditional on correct report of T1 (T2|T1), T2|T1 accuracy if T1 is followed by a blank (T1+1) reproduced from Chun and Potter (1995). T2|T1 accuracy if T2 is placed at the end of the RSVP stream reproduced from Giesbrecht and Di Lollo (1998).

Unless target representations are fleeting, observers seldom fail to detect a stimulus in RSVP. Hence, the AB occurs only if the second target is masked (Giesbrecht, Bischof, & Kingstone, 2003; Grandison, Ghirardelli, & Egeth, 1997; Seiffert & DiLollo, 1997; Dell'Acqua, Pascali, Jolicoeur, & Sessa, 2003). If T2 is unmasked by placing it at the end of the RSVP stream (Figure 9), detection performance is at ceiling throughout the AB (Giesbrecht & Di Lollo, 1998; Vogel & Luck, 2002) .

If T1 is followed by a blank interval, this reduces masking effects on T1, which leads to a shallower and shorter AB (Figure 9, see also Chun & Potter, 1995). The T1+1 blank finding is troublesome for pure competition based theories of the AB, as they would predict that a stronger T1 suppresses T2 to a greater extent and this should produce a deeper AB. It thus seems that one of the critical factors influencing the AB is the amount of time it takes to consolidate T1 into working memory. Hence, there is a reciprocal relationship between T1 strength and the extent of the AB (Bowman et al., 2008). Removing the T1 forward mask (T1-1 blank) has no effect on the AB (Breitmeyer, Ehrenstein, Pritchard, Hiscock, & Crisan, 1999). It thus seems that the AB is mainly dependent on backward and not forward masking effects (see Chapter 5 for an extensive discussion of this issue).

Traditionally, it was thought that 'T1 processing must be interrupted by another visual stimulus or the AB effect will not occur' (Shapiro, Raymond, & Arnell, 1997). A recent experiment, however, challenges this hypothesis by finding that even if T1 is completely unmasked the AB still occurs, providing that T2's representation is rendered weak enough (Nieuwenstein, Potter, & Theeuwes, 2009). The authors find that if T2 is presented for only 58ms (instead of the typical 100ms) and followed by a powerful pattern mask, participants show a strong AB even if there are no distractors between T1 and T2.

### The attentional blink and conscious perception

Sergent and Dehaene (2004) employ an AB paradigm where participants are asked to rate the visibility of T2 on a continuous scale. The behavioural results suggest that participants use the scale in an all-or-none fashion for T2s during the AB, i.e. if T2 is presented at lag 3 following an identified T1 (Figure 10A). If observers are instructed to ignore T1 and just report T2 (i.e. the equivalent of T2 being presented outside the AB), they respond in a gradual manner (Figure 10B).

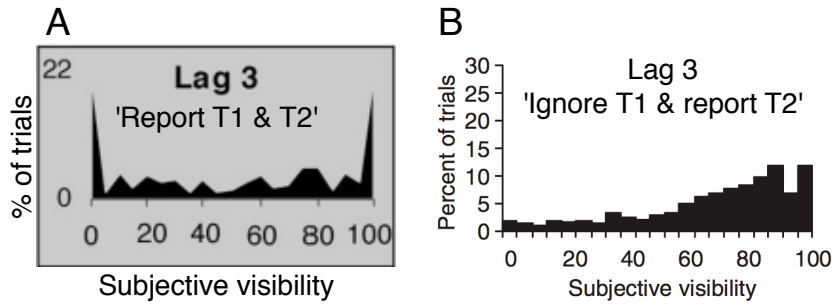


*Figure 10* Panel A: Response distribution (percentage of trials per visibility score) for a T2 presented during the AB, i.e. T2 follows T1 at lag 3 and participants are instructed to report both T1 and T2. Adapted from Sergent and Dehaene (2004). Panel B: Response distribution (percentage of trials per visibility score) for a T2 presented outside the AB, i.e. T2 follows T1 at lag 3 but participants are instructed to ignore T1. Adapted from Sergent et al. (2005).

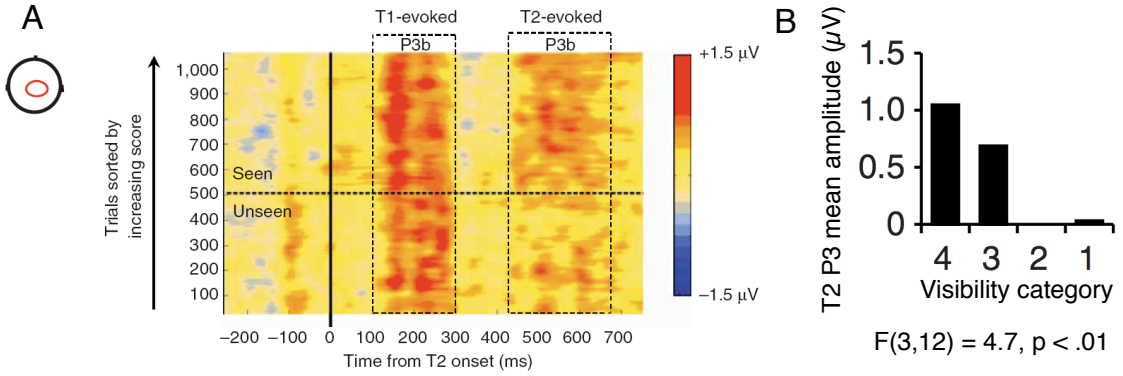


*Figure 11* Panel A: ERPimage plot of the P3 components evoked by T1 and T2 when T2 is presented during the AB (at lag 3). Trials are sorted by visibility score, lowest visibility at the bottom. Voltage values (positive in red and negative in blue) are smoothed over 50 trial windows. Panel B: Histogram of mean amplitude of the T2 P3 component (presented during the AB) per visibility category. Categories 4 and 3 correspond to visibility scores $> 50\%$ ('seen' trials), whereas categories 2 and 1 to visibility scores $< 50\%$ ('unseen' trials). Within seen and unseen trials, the categories are classified per participant using the median of that participant's response distribution. Adapted from Sergent et al. (2005).

In a follow-on study to their behavioural analysis, Sergent et al. (2005) conduct an EEG study using the same paradigm. The authors present an ERPimage showing the trial-by-trial P3 component for targets sorted by their visibility score (Figure 11A). There is an effect of visibility score on P3 amplitude and the histogram indicates a difference between P3 size for high compared to low visibility categories (Figure 11B). Based on these findings, Sergent et al. (2005) conclude that in addition to behavioural report being all-or-none (Sergent & Dehaene, 2004), the distribution of P3 component sizes for targets during the AB is also bimodal. They suggest that conscious perception during the AB is all-or-none, i.e. people either perceive the target or completely miss it.

Correctly identified stimuli in RSVP have been found to correlate with increased EEG oscillations in the gamma frequency band (Kranczioch et al., 2006; Nakatani, Ito, Nikolaev, Gong, & van Leeuwen, 2005), which suggests a relationship between EEG oscillations in the gamma band and conscious perception. Accordingly, targets that are seen during the AB produce increased power in the gamma band of EEG oscillations compared to targets that are missed during the AB (Fell, Klaver, Elger, & Fernández, 2002; Kranczioch, Debener, Maye, & Engel, 2007).

### Locating the attentional blink

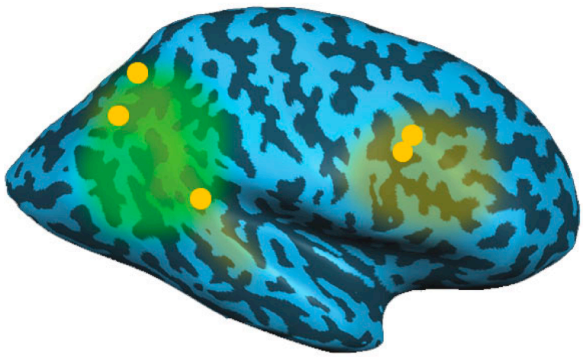

*Figure 12* Neuroanatomical substrates of the AB. The gold circles represent activation foci from fMRI studies, whereas the light- and dark-green shaded regions represent areas implicated in lesion and MEG studies, respectively. Adapted from Marois (2005).

A wide range of studies have explored the neural correlates of the AB. Lesion studies provide evidence pointing towards the right brain hemisphere being the source of the

AB (Giesbrecht, Bischof, & Kingstone, 2004; Hillstrom, Husain, Shapiro, & Rorden, 2004). In addition, functional Magnetic Resonance Imaging (fMRI) research suggests stronger activation in lateral frontal and parietal areas for seen compared to missed T2s during the AB (Kranczioch, Debener, Schwarzbach, Goebel, & Engel, 2005; Marois, Chun, & Gore, 2000; Marois, Yi, & Chun, 2004). Synchronised oscillations in the beta band between the neural populations in these two relatively distant regions (frontal and parietal) might play an important role for attentional processes required during the AB (Gross et al., 2004; Kranczioch et al., 2007).

The evidence from fMRI and magnetoencephalography (MEG) studies suggests that the neural substrates of the AB lie in frontal and parietal brain regions (Figure 12). Transcranial Magnetic Stimulation (TMS) directed at the right Intraparietal Sulcus (IPS) area seems to decrease the AB impairment both in terms of length (Kihara et al., 2007) and magnitude (A. Cooper, Humphreys, Hulleman, Praamstra, & Georgeson, 2004), whereas stimulation of a control condition (central and parietal midline electrode position) has no effect on the AB. Hence, in parietal regions, the right IPS seems to be one of the neural correlates of the AB.

# Part II

# Applying cognitive electrophysiology to neural modelling of the attentional blink

# Chapter 3

# Theories of the attentional blink

This thesis presents experimental data that are used to assess current theories of the AB. In order to interpret the implications of our results, the following chapter reviews current theories of the AB. Section 3.1 introduces the most influential informal theories of the AB. These theories are informal in the sense that they have a textual description but no mathematical implementation that could be formally tested.

Section 3.2 describes current formal models of the AB. These models are inspired by informal theories of the AB. However, in addition to providing a detailed description of the theory, the formal models simulate behavioural output, which can be interpreted to validate the underlying theory. For those models that are implemented using neural network architectures, we assess how well their neural activation traces match the profile of the human EEG. Hence, we outline the current state of the art in generating virtual ERPs from neural network models of the AB in order to put the methodological contribution of this thesis (i.e. generating virtual ERP from the $ST^2$ model; see Chapter 4) into context. Finally, this chapter concludes with a detailed description of the $ST^2$ model in Section 3.3.

## 3.1 Informal theories of the attentional blink

### 3.1.1 Distractor induced suppression

One of the first theories of the AB proposed that it is the distractor in the T1+1 position that causes the impairment in the detection of T2 (Raymond et al., 1992). This

distractor induced suppression theory has recently been extended by the boost and bounce model (Olivers & Meeter, 2008), which is a formal model of the AB and is discussed in Section 3.2.4. According to the distractor induced suppression theory, T1 opens up an attentional gate, which is 'sluggish' to close. Due to the fast presentation rate of RSVP, the attentional gate is often still open when the T1+1 item is presented, hence, this item can also enter later stages of processing. If the T1+1 item is another target, there is a high probability of that target being seen, as exemplified by lag 1 sparing. However, if the T1+1 item is a distractor, this causes a disruption to the task set. The disruption enables a period of suppression that can last for several hundred milliseconds. The system cannot process further targets during the suppression period, which results in the reduction of T2 accuracy that is observed during the AB. Targets can only be accurately identified during the AB if they are strong enough to survive this period of suppression. In addition to explaining the traditional AB curve, this theory can also account for spreading the sparing (Olivers et al., 2007), as the theory predicts high performance on target detection as long as the task set is not disrupted by a distractor. However, the distractor induced suppression account has great trouble explaining the finding of a strong AB despite T1 being unmasked, i.e. if the distractors following T1 are omitted (Nieuwenstein et al., 2009). If the AB is caused by the distractors following T1, as this theory proposes, why does the AB still occur even if T1 is unmasked?

### 3.1.2 Two-stage theory

This theory proposes two stages of visual processing in RSVP (Chun & Potter, 1995). All items presented in the RSVP stream are processed for sensory and limited semantic features in a first stage. Since the first stage is parallel, more than one item representation can be active at a time. If an item matches the task template, it is categorised as a probable target and triggers a transient attentional response. The attentional response enhances processing of the target (Nakayama & Mackeben, 1989; Weichselgartner & Sperling, 1987) and allows the target to proceed to a second stage. However, as the second stage can only process one item at a time, the system encounters a bottleneck. Subsequent targets have to wait at stage one while stage two is occupied with encoding the previous target into working memory. In the AB context, stage two is occupied for approx. 200-600ms following T1

presentation while encoding T1. As the T2 is subject to decay while waiting for encoding resources to free up, T2 is more likely to be missed during this period, which corresponds to the reduction in T2 performance during the AB.

Since the attentional response outlasts T1 presentation, both T1 and its following item in the RSVP stream are enhanced. The two stage theory thus proposes that T1 and the following item often enter stage two together (see Chapter 6 for a further discussion of this issue). If T2 is presented at lag 1, it will benefit from this joint processing and, indeed, T2 accuracy is often at or even above baseline performance from outside the AB (lag 1 sparing). However, joint encoding at lag 1 leads to a loss of order information (i.e. increased number of swaps) and trade-off effects (see also Section 2.3.2).

### 3.1.3 Interference theory and resource sharing

The interference theory (Shapiro, Raymond, & Arnell, 1994, 1997) and two-stage accounts both propose that all items in the RSVP stream are processed to a sensory and limited semantic level. However, the theories diverge in explaining how target items are encoded into working memory during the AB. According to the interference theory, target items (T1 and T2) match a predefined template or task filter, which means that they are considered for further processing. The distractors (masks) following T1 and T2 appear in such close temporal proximity to the target that they are also consolidated into visual short term memory (VSTM). In VSTM, the two targets and their masks compete throughout the duration of the AB based on their strength values, which are determined by a weight assigned to each of the items. T1 receives a stronger weighting as it is presented first, whereas T2, as the second task, has a lower weight. Due to the competition/interference during the AB, the lower weight value assigned to T2 results in T2 often being missed during the AB.

In recent articles, the interference theory seems to have been superseded by the resource sharing hypothesis. The resource sharing hypothesis suggests that there is a limited amount of shared resource available for processing targets during the AB (Shapiro, Schmitz, Martens, Hommel, & Schnitzler, 2006). Although there is competition between the targets throughout the duration of the AB, unlike in the interference theory the competition is not based on bottom-up target strength but rather on the amount of resource that people -

whether it be voluntary or involuntary - allocate to T1 and T2. Although not specifically stated by the authors, it can be assumed that, as proposed by the interference theory, all targets are processed at a sensory and also some semantic level. Resource allocation then has its effect once the targets have been consolidated into VSTM. If less resource is allocated to T1, more resource is available for T2 and, vice versa, if more resource is allocated to T1, less resource is available for T2 (Shore, McLaughlin, & Klein, 2001; Kranczioch et al., 2007). Thus, the AB is caused by over-investment of resource in T1 processing, which means that there is not enough resource available to process T2 and it is often missed. In turn, if people succeed in allocating less resource to T1 processing (either voluntarily (Martens, Munneke, et al., 2006; Slagter et al., 2007) or involuntarily (Olivers & Nieuwenhuis, 2005; Ho et al., 2007)) more resource is available to process T2, which increases T2 accuracy. See Chapter 6 for an extensive discussion of the resource sharing theory.

The *spreading the sparing* (Olivers et al., 2007) and *whole report* (Nieuwenstein & Potter, 2006) findings are troublesome for competition-based accounts that propose a general notion of limited cognitive resources during the AB (Kranczioch et al., 2007; Shapiro et al., 2006; Shapiro, Raymond, & Arnell, 1997). In these experiments, trials with varying number of targets appear at random. Hence, participants are not aware of the type of trial (i.e. the number of targets that a trial will contain) before the trial has been presented, but are nevertheless able to 'spread the sparing'. If there was a fixed amount of cognitive resource that could be distributed between the targets, it is unclear how people would know beforehand how much resource to allocate to each of the targets. Alternatively, one might argue that the resource allocation occurs instantaneously. However, this assumption imposes the question of how resources could be 're-allocated' to new targets despite having already been deployed to process a preceding target. Furthermore, why is average performance across all targets higher when four targets are presented in a row (i.e. during whole report) than when only two targets are presented (as is the case in the classical AB paradigm)? In terms of the resource sharing hypothesis this would suggest that the absolute amount of resource suddenly increases when a sequence of four targets is presented compared to when a trial contains only two targets.

### 3.1.4 Temporary loss of control

In similar fashion to the distractor induced suppression theory (Raymond et al., 1992), the temporary loss of control hypothesis (TLC; Kawahara et al., 2006; Di Lollo et al., 2005) proposes that the distractor in the position following T1 is the cause of the AB. Specifically, the TLC theory suggests that the T1+1 distractor causes a temporary loss of control over the stimulus input. Prior to the occurrence of T1, the visual system's input filter is configured to let targets pass and suppress distractors. However, while T1 is being consolidated, the system can no longer maintain that attentional set and the input filter becomes vulnerable to stimulus-driven disruption from distractors. Consequently, if a distractor is presented in the T1+1 position, this causes the input filter to be erroneously reconfigured. Hence, a following T2 does not match the input filter and is often missed, which causes the AB. If there are no intervening distractors, on the other hand, the system is able to detect multiple targets in a row (as seen in lag 1 sparing (Chun & Potter, 1995), spreading the sparing (Olivers et al., 2007) or whole report (Nieuwenstein & Potter, 2006)). However, as with the distractor induced suppression and boost and bounce theory (Raymond et al., 1992; Olivers & Meeter, 2008), the TLC theory has trouble explaining the result that observers show a strong AB even if T1 is unmasked (Nieuwenstein et al., 2009).

## 3.2 Formal models of the attentional blink

In the following section, we discuss those formal models of the AB that (a) build upon a neural network architecture and (b) are concerned with the letters-in-digits AB task. For each of these models, we first describe their architecture, then how the model simulates the AB and, finally, evaluate the neural activation patterns that can be extracted from the neural network implementation.

Other formal models, such as the computational implementation of the Interacting Cognitive Subsystems (ICS) model by Barnard and Bowman (2004), the conflict-monitoring model by Battye (2003) or the auto-associator model by Chartier, Cousineau, and Charbonneau (2004), are not discussed. These models are either concerned with different tasks (targets marked by semantics or colour) and/or are not implemented using a neural network. A detailed review of these other formal models of the AB can be found in Bowman

and Wyble (2007).

### 3.2.1 Global workspace model

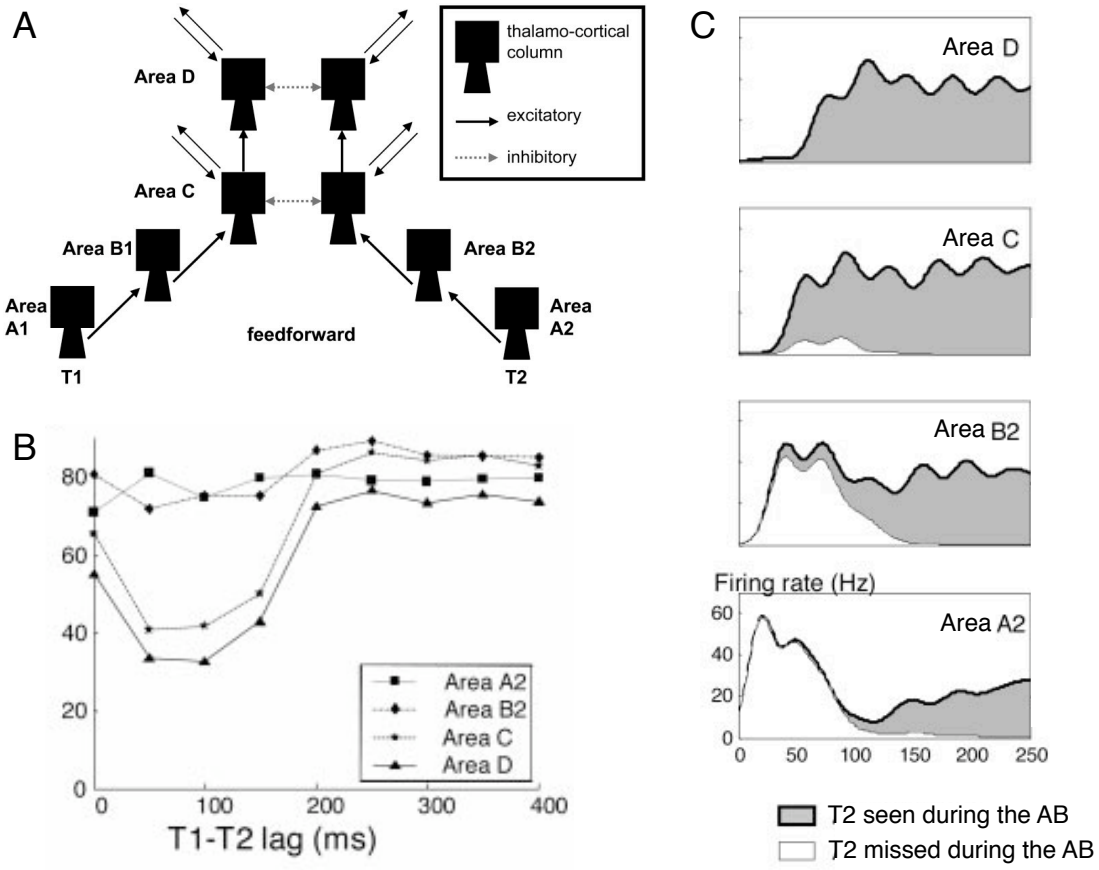


*Figure 13* Panel A: Processing pathways in the global workspace model. Panel B: Areas C and D represent conscious perception, hence, the corresponding curves of the figure reflect the simulated behavioural accuracy of the global workspace model during the AB. Panel C: Neural activity evoked by seen and missed targets during the AB. Areas A2 and B2 correspond to early sensory processing, whereas areas C and D correspond to later stages of processing. Adapted from Dehaene et al. (2003).

The global workspace model (Dehaene et al., 2003) describes neural processing pathways from early sensory up to higher processing (Figure 13A). Stimuli compete to enter a global workspace, which allows stimuli to be consciously perceived. A unitary nature of consciousness is achieved as the neurons representing a stimulus in the global workspace

inhibit neurons of surrounding items. Representations are sustained through reverberating activation feeding back to initial perceptual stages. The global workspace model's strength is its close tie with biology as the authors use spiking neurons in their neural network implementation and connect parts of the model to specific brain areas.

#### How the global workspace model blinks

Figure 13A illustrates how the global workspace model blinks. T1 and T2 are processed along separate neural processing pathways. T1 is presented first and enters the global workspace (or consciousness). The firing rates of neurons at the top of the processing pathway (areas C and D in Figure 13A) indicate whether a target was seen or missed. If T2 is presented during the AB, T1 is being processed in the global workspace and inhibits T2, which decreases T2 accuracy. Once T1 processing has completed, T2 is more likely to be reported, which corresponds to the period after the AB.

#### Neural activation patterns from the global workspace model

Areas A2 and B2 are associated with sensory and semantic processing, whereas C and D are supposed to represent higher level processing. As illustrated in Figure 13C, areas A2 and B2 remain active also for T2s that are missed during the AB. Areas C and D, however, are active only for T2s that are seen but not for targets that are missed during the AB. Thus, the firing patterns from the global workspace model show a qualitative match of the ERP effects observed during the AB (see Section 2.3.2).

#### Evaluation

The global workspace model produces reduced T2 accuracy if T2 is presented within 200ms following T1 (Figure 13 B). Although the simulated behavioural accuracy curve looks similar to an AB curve, the simulated behavioural data show a number of discrepancies in relation to the human behavioural data during the AB. First, the global workspace model's AB curve is too short by about 400ms, as the AB typically lasts for around 600ms. Furthermore, at lag 1, where one would typically expect lag 1 sparing, the model produces the lowest T2 performance whereas T2 performance is high at lag 0. However, at lag 0 T1 and T2 would

be presented simultaneously, which normally cannot happen in an AB paradigm due to the serial nature of RSVP.

The global workspace model can be connected to electrophysiology by means of neural firing patterns. The timing of the activation dynamics, however, does not match the data from human experiments, which might be due to restrictions enforced by the use of parameters from monkey neurophysiology. Hence, the neural firing patterns from the global workspace model replicate some of the ERP effects observed during the AB, but the connection remains of a loose qualitative nature.

### 3.2.2 The CODAM model

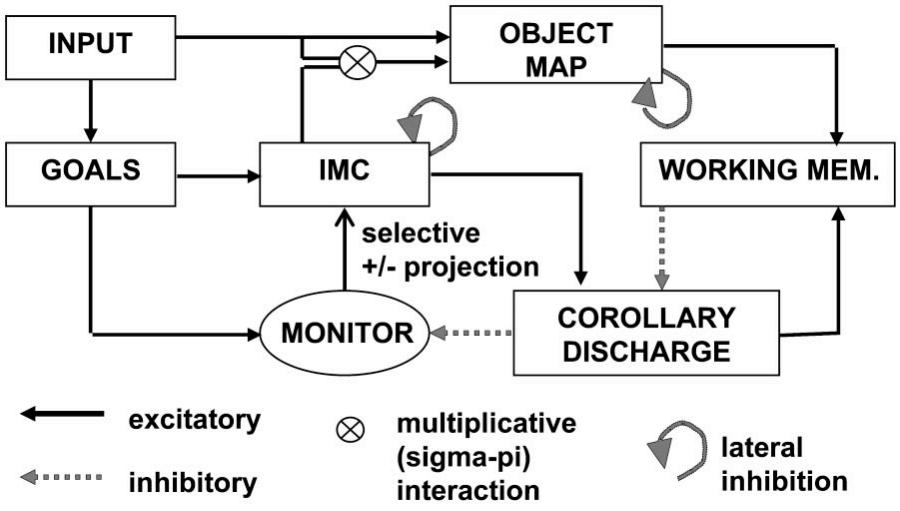


*Figure 14* The version of the CODAM model used to simulate the attentional blink. Adapted from Fragopanagos et al. (2005).

Fragopanagos et al. (2005) model the AB in the context of the corollary discharge of attention movement model (CODAM; Taylor & Rogers, 2002). CODAM is a neural network model and consists of modules that interact via excitatory and inhibitory connections (Figure 14). Stimuli are processed in the input and object map modules before progressing to the working memory module, where items are encoded into working memory. For items to reach the working memory layer, however, they require an attentional enhancement from the inverse model controller (IMC).

### How the CODAM model blinks

The IMC module uses the monitor, goals and corollary discharge modules to constantly compare items presented in the RSVP stream to the current target template. If items match the template, they receive an attentional enhancement and can be encoded by the working memory module. In the context of the AB, T1 is the current target for the IMC until it has been encoded. For the duration of the AB, T2 thus does not match the IMC target template and consequently does not receive an enhancement from the IMC. In consequence, T2 often does not gain sufficient activation to progress to the working memory module, which results in the reduction in T2 accuracy during the AB.

### Neural activation patterns from the CODAM model

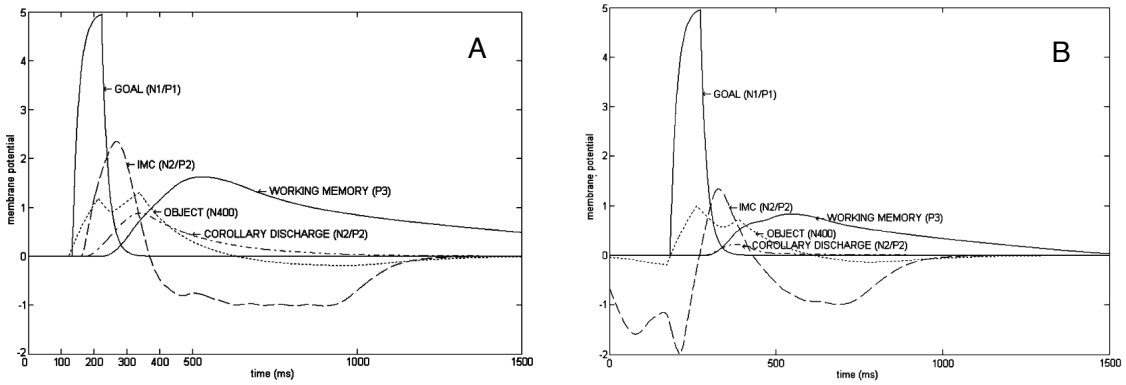


*Figure 15* Panel A: Membrane Potentials from CODAM for T1. Panel B: Membrane Potentials from CODAM for T2 presented during the AB. Adapted from Fragopanagos et al. (2005).

Apart from simulating behavioural results, Fragopanagos et al. (2005) present a qualitative match between membrane potentials from CODAM neural network nodes and ERP components observed during the AB. Figure 15A shows the CODAM module activation patterns for T1, whereas Figure 15B illustrates the CODAM activation profiles for T2s presented during the AB. The CODAM modules connected to sensory and semantic processing remain active for T2s during the AB, thus qualitatively matching the human ERP data (see Section 2.3.2). The activation trace of the working memory module, however, is reduced for T2s during the AB in comparison to the working memory module activation for T1. As the working memory module activation is associated with the P3 component,

this provides a qualitative replication of the P3 effect observed during ERP experiments of the AB (see Section 2.3.2).

#### Evaluation

The CODAM approach makes a valuable contribution as it links a neural network, which is capable of simulating behavioural results for a variety of paradigms (not just the AB), to specific ERP components. It also provides an indication as to where parts of the CODAM model lie in the visual processing pathway. Furthermore, the membrane potentials from CODAM provide a qualitative fit for most of the ERP effects observed during the AB (Vogel et al., 1998). Inconsistencies, however, occur in terms of the relationship between the activation profiles for the individual CODAM modules and their associated human ERP components. Figure 15 shows that the object model trace, for instance, occurs around 400ms earlier than the working memory trace. In the human ERP, however, the associated ERP components appear in the opposite order, i.e. the P3 usually precedes the N400 component. Furthermore, traces associated with early components have larger amplitudes than traces associated with the P3. In the human ERP this relationship is reversed.

### 3.2.3 The LC-NE model

The locus coeruleus-norepinephrine (LC-NE) model (Nieuwenhuis, Gilzenrat, et al., 2005) proposes a neurophysiological basis for the AB. It has been suggested that the minute brain-stem structure locus coeruleus (LC) is critical for the regulation of cognitive performance through the release of the neuromodulator norepinephrine (NE) to widespread cortical areas (Aston-Jones, Rajkowski, & Cohen, 2000). As shown in Figure 16, the LC-NE model consists of three layers that form the behavioural network and the LC module. Although all items feed through the behavioural network, only targets can initiate the LC response, which is required for targets to enter the detection layer and be encoded into working memory.

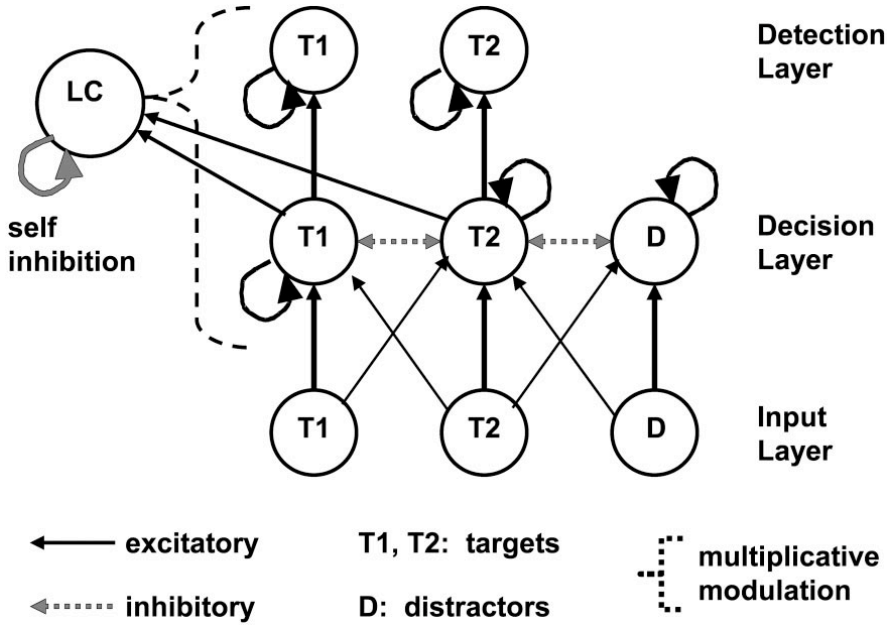


*Figure 16* Depiction of the locus coeruleus - norepinephrine (LC-NE) model. Adapted from Nieuwenhuis, Gilzenrat, et al. (2005).

### How the LC-NE model blinks

The refractory period of the LC is what causes the AB. After having fired for T1, the LC is in its refractory period if T2 is presented during the AB. T2 cannot be enhanced by the LC and often fails to progress to the detection layer, which accounts for reduced T2 accuracy during the AB. The LC-NE model replicates a basic U-shaped AB curve, where 'lag 1 sparing' occurs because T2 receives the benefit of the LC response initiated by T1.

### Neural activation patterns from the LC-NE model

The P3 has been proposed to reflect phasic activity of the LC-NE system (Nieuwenhuis, Aston-Jones, & Cohen, 2005). Figure 17 shows how the artificial LC response generated by the LC-NE model is reduced for T2s that are missed during the AB compared to T2s that are seen, which (if the LC-NE response is indeed related to the P3) is in line with results from ERP studies investigating the AB.

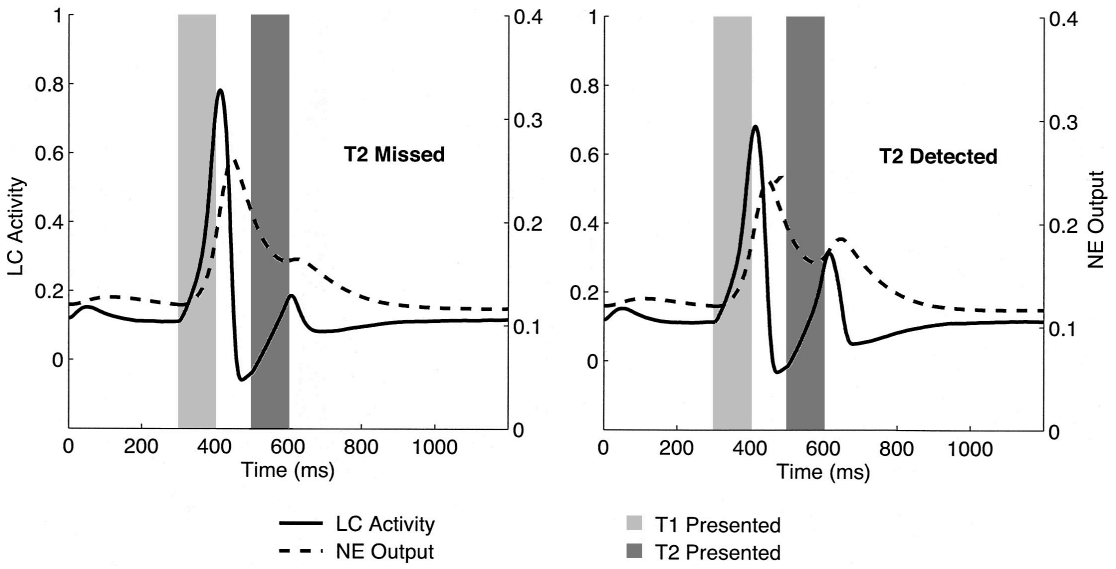


*Figure 17* Activation dynamics of the abstracted locus coeruleus (LC) for missed and seen T2s during the AB. LC activity and norepinephrine (NE) output are scaled on separate axes. Time 0 ms indicates the onset of the simulated trials. Adapted from Nieuwenhuis, Gilzenrat, et al. (2005).

### Evaluation

According to the LC-NE theory, there is a direct correlation between the neuromodulator NE (released by the LC) and the impairment observed during the AB. A recent pharmacological study (Nieuwenhuis, van Nieuwpoort, Veltman, & Drent, 2007) tested this hypothesis using the $\alpha 2$ adrenoceptor agonist clonidine, which at low doses decreases LC firing and attenuates the release of NE from axon terminals (Svensson, Bunney, & Aghajanian, 1975). Participants performed an AB task and a visual search task and were split into two groups. One group received an oral dose of clonidine and the other received a placebo. Whereas clonidine slowed overall reaction times in the visual search task, no effect was found for the AB task. Thus, these results speak against an involvement of NE on subjects' performance during the AB and indeed provide strong evidence against the theory underlying the LC-NE model.

## 3.2.4 The boost and bounce model

The boost and bounce model (Olivers & Meeter, 2008) further develops the distractor induced suppression theory originally proposed by Raymond et al. (1992) (see Section 3.1.1)

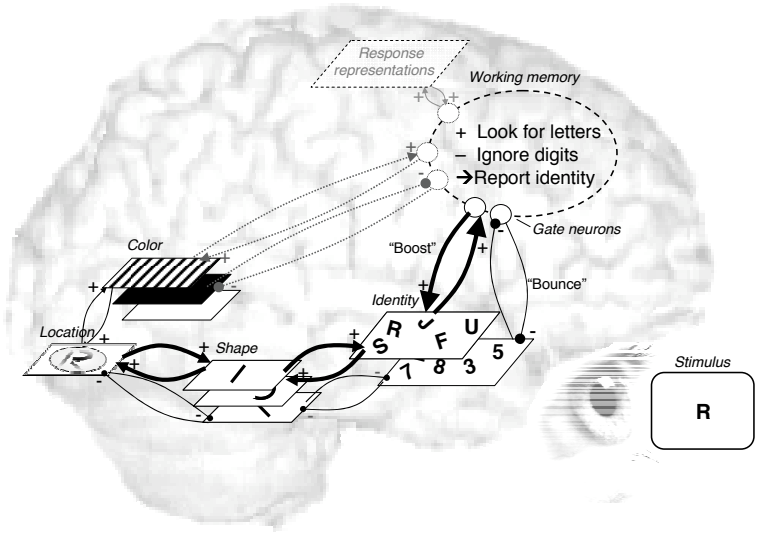


*Figure 18* Schematic depiction of the boost and bounce model. As stimuli representations feed through the model, they are processed for colour, shape and identity. The task the observer is asked to do (report the letter among the digits in the example shown here) determines the attentional set, which is implemented in the working memory gating system. The gating system works by selectively enhancing ('boosting') targets and inhibiting ('bouncing') distractors and thus allowing targets to enter working memory. Adapted from Olivers and Meeter (2008).

in that the boost and bounce model hypothesises the distractor in the T1+1 position to be the cause of the AB. The boost and bounce model's main theoretical distinction in comparison to other formal models of the AB is that there is 'no central role for capacity limitations or bottlenecks' (Olivers & Meeter, 2008). Rather, visual perception is controlled by a rapidly responding attentional filter (or gate) that enhances ('boosts') relevant and suppresses ('bounces') irrelevant information. As shown in Figure 18, the selective enhancement and inhibition is implemented using excitatory and inhibitory connections.

### How the boost and bounce model blinks

According to the boost and bounce model, the AB is caused by a malfunction of the attentional gating system. Until T1 is presented, the attentional gate is open and configured to let targets pass. If, however, T1 is followed by a distractor, this distractor receives a large amount of excitatory feedback, but does not match the target template. The distractor thus triggers a strong inhibitory response from the neurons in the attentional gate. This prevents the distractor from being misperceived as a target, but also initiates a prolonged period of

suppression that prevents any following targets from entering working memory and results in the AB.

The boost and bounce model can account for people not showing an AB if multiple targets are presented in a row (i.e. lag 1 sparing, spreading the sparing or whole report; see Section 2.3.2). In these paradigms, the first target is not followed by a distractor but instead followed by another target, which is of the same category. Consequently, there is no disruption to the task set and the attentional gate continues to let targets pass and to suppress distractors. In these circumstances, target perception is limited by the number of items that can be simultaneously stored in working memory.

### Neural activation patterns from the boost and bounce model

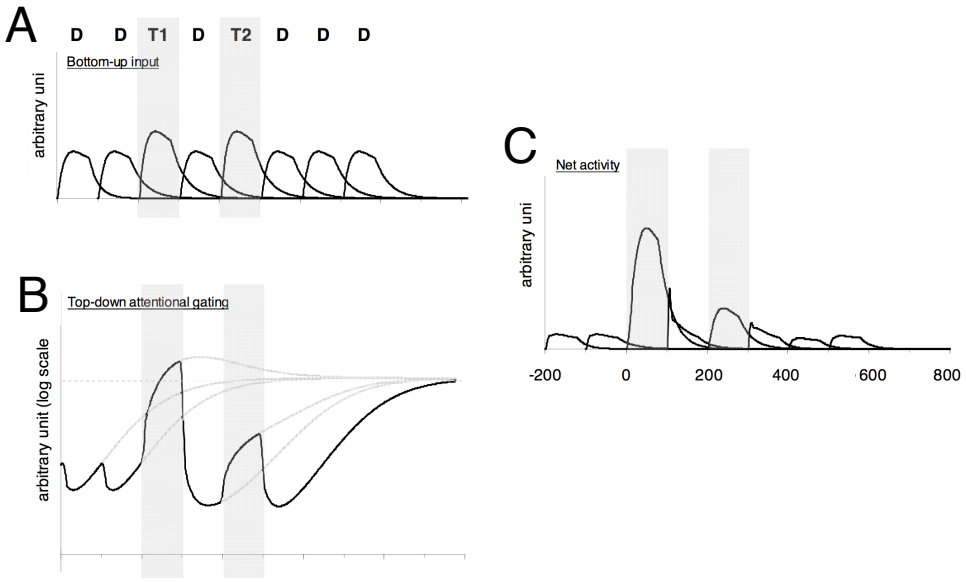


*Figure 19* Neural dynamics of the boost and bounce model for a target presented during the AB (T2 presented at lag 2). Panel A: Activation from nodes resembling bottom-up sensory processing. Panel B: Top-down attentional response to the input (combination of both excitatory and inhibitory feedback). Panel C: Combined activation trace of bottom-up and top-down signal. Adapted from Olivers and Meeter (2008).

Figure 19 shows neural activation traces from the boost and bounce model. As seen in Figure 19A, each item presented in the RSVP stream evokes a neural response whereby

the two targets (T1 and T2) have more bottom-up signal than distractors. The behaviour of the attentional gate is demonstrated in Figure 19B. Although the presentation of T1 causes excitatory activation, the distractor following T1 leads to a inhibitory response. The following disruption of the system means that T2's excitatory response is smaller with the consequence being the AB. The interplay of bottom-up signal and top-down attentional response is also illustrated in Figure 19C, which shows the combined activation trace.

#### Evaluation

The boost and bounce model provides an elegant explanation for both the inability to detect targets during the AB and *spreading the sparing* (see Section 2.3.2); two findings that, at first sight, seem to contradict each other. As it has a computational neural network implementation, the boost and bounce model can be used to run simulations and generate predictions for experimental data.

Although the authors provide an indication as to where parts of the model lie in terms of the visual processing hierarchy in the brain (Figure 18), they do not associate the activation traces from their model with specific ERP components. One could speculate, however, that the bottom-up input signal from Figure 19A represents a notion of early components (like the ssVEP wave; see Section 2.1.1). The combined activation trace in Figure 19C might be associated with the P3 component (see Section 2.1.1), which would qualitatively replicate the finding of a smaller P3 for T2 than for T1. There is a discrepancy in terms of latency, however, as the model's activation trace starts almost immediately after the target has been presented (Figure 19C), whereas the P3 component onsets with a latency of several hundred milliseconds. A further mismatch concerns the duration of the model's combined activation trace, as it only lasts for 100ms, which, in fact, corresponds to the presentation time of stimuli.

## 3.3 The $ST^2$ model

In the following section, we explain the fundamental principles of how the $ST^2$ model (Bowman & Wyble, 2007) simulates working memory, temporal attention and, in particular, the attentional blink. The $ST^2$ model is described to the extent required for the work

presented in this thesis; please refer to Bowman and Wyble (2007; pages 41-51 and 68-69) for a full description of the $ST^2$ model and the mathematical details of its neural network implementation ('neural-$ST^2$').

### 3.3.1 Types & tokens

The $ST^2$ model employs a types-tokens account (Kanwisher, 1987; Chun, 1997b) to describe the process of working memory encoding. Types describe all feature related properties associated with an item. These include sensory properties, such as visual features (e.g. its shape, colour and the line segments comprising it) and also semantic attributes, such as a letter's position in the alphabet. A token, on the other hand, represents episodic information, which is specific to a particular occurrence of an item, thus providing a notion of serial order. An item is encoded into working memory by creating a connection between a type and a token. At retrieval, tokens contain information about when an item occurred and, from tokens, types can be regenerated, yielding a description of what each item was and in what temporal order they appeared.

### 3.3.2 Model architecture

As illustrated in Figure 20, the $ST^2$ model can be divided into three parts. We describe them in turn:

**(1) Stage 1: Input & extraction of types** The extraction of types in stage one occurs in four layers. Input values, which simulate target letters and digit distractors in the AB experiment, are fed into the input layer of the model. Accordingly, at any timestep, this layer represents the stimulus currently being presented to the model. As activation values propagate upwards, the following layers (masking, item and task filtered layer) reflect forward and backward masking in early visual processing and the extraction of semantic representations. A task demand mechanism operates at the task filtered layer (TFL) and ensures that only targets are selected for working memory encoding. Despite the fact that stimuli are presented serially during the AB task, processing within stage one may exceed the presentation time of sequentially presented items. Hence, these layers are parallel or simultaneous in nature, in that more than one node can be active at a time.

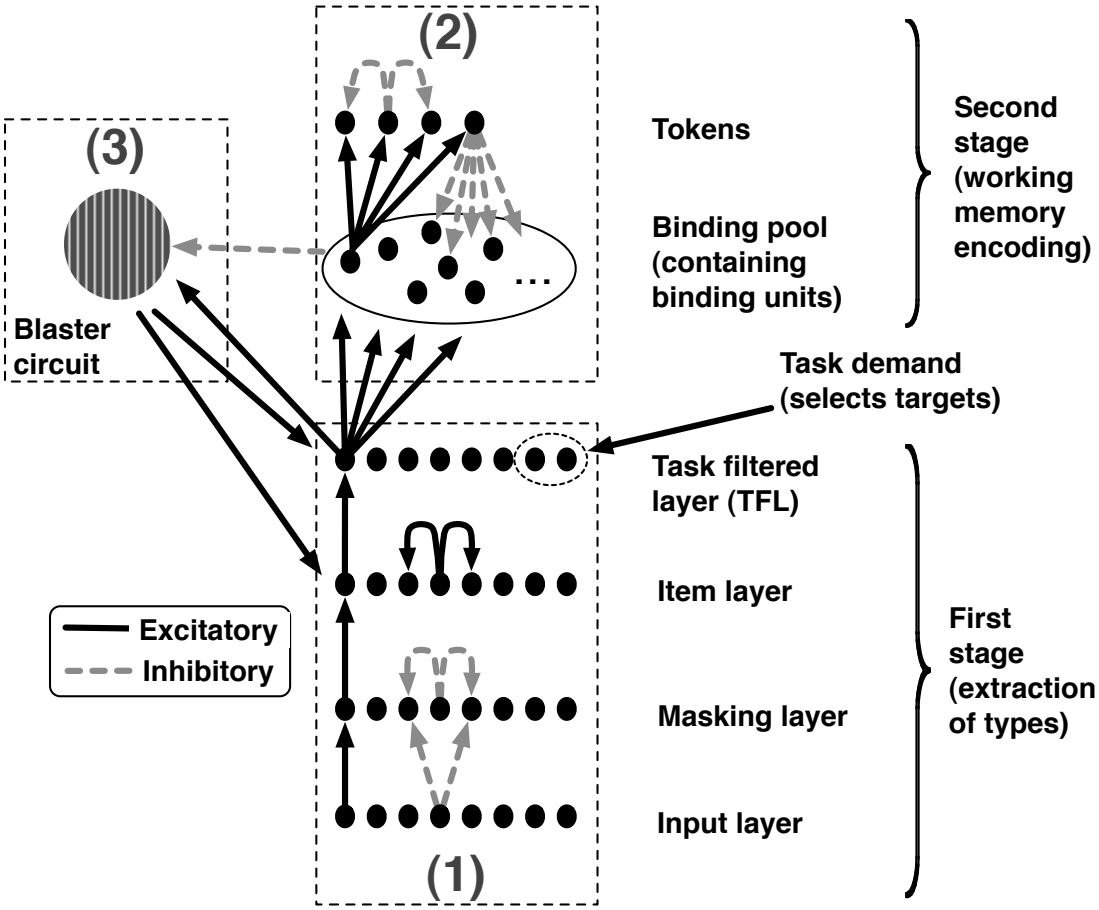


*Figure 20* The $ST^2$ model: (1) Input & extraction of types in stage one, (2) Working memory tokens in stage two, (3) Temporal attention from the blaster. Adapted from Bowman and Wyble (2007).

**(2) Stage 2: Working memory encoding** An item is encoded into working memory by connecting its type from stage one to a working memory token from stage two. This process is referred to as 'tokenisation' (for further details please refer to Bowman and Wyble (2007)). If at the end of a trial, the type node of a target has a valid connection to a token, the target is successfully 'reported' by the $ST^2$ model. Inhibition between working memory tokens ensures only one token is active at a time. This means that only one tokenisation process can be active at a time, thus enforcing a serial nature of working memory encoding.

**(3) Temporal attention from the blaster** Temporal attention is implemented by a mechanism termed the blaster. Salient items in the TFL trigger the blaster, which provides a powerful enhancement to all nodes of the item layer and the TFL. The enhancement from the blaster allows targets to become sufficiently active to initiate the tokenisation process. During tokenisation, the blaster is suppressed until encoding of the target has completed. The suppression prevents a second target from refiring the blaster while the first target is being tokenised, which would corrupt the working memory encoding process.

### 3.3.3 How the $ST^2$ model blinks

During the AB, T1 is in the process of being tokenised when T2 is presented, thus, the blaster cannot enhance T2 as the blaster is suppressed during the tokenisation of T1. By the time T1 tokenisation has completed, T2 will often lack sufficient activation to initiate its own tokenisation process, which causes T2 to be missed, resulting in an AB. The duration of the AB thus corresponds to the amount of time it takes to tokenise T1, which, in turn, depends on its relative trace strength. If, however, a T2 item is particularly salient (e.g. the participant's name (Shapiro, Caldwell, & Sorensen, 1997)) or subject to less masking (e.g. T2s at the end of the RSVP stream (Giesbrecht & Di Lollo, 1998; Vogel & Luck, 2002)) it often has sufficient trace strength to 'outlive' T1's tokenisation. Such T2s will be seen during the AB.

At lag 1, however, T2 is presented during the window of blaster enhancement initiated by T1. T2 gains sufficient activation to join T1's working memory encoding process and T1 and T2 are tokenised together. This leads to increased T2 performance, as exemplified by lag 1 sparing. However, T1 and T2 are swapped more frequently than warranted by chance

as there is no notion of serial order within a token (Bowman & Wyble, 2007). Particularly strong T2s can suppress T1 during the process of joint working memory encoding and lead to reduced T1 performance (see Section 2.3.2), but this only obtains at lag 1.

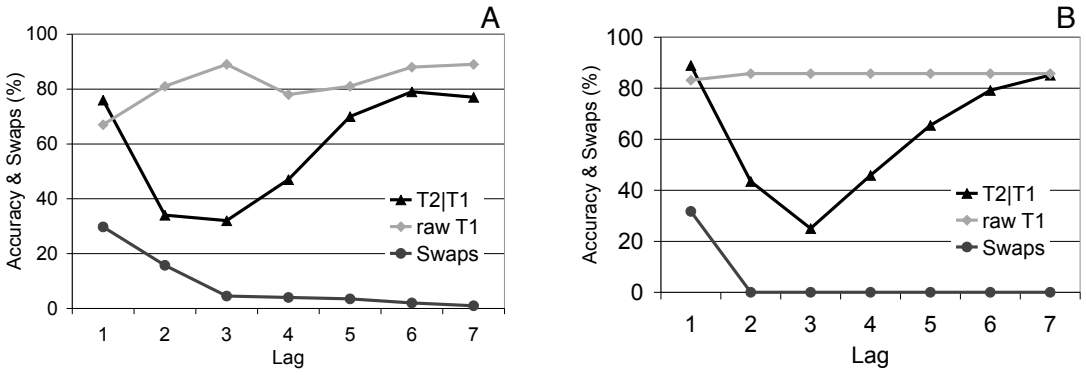


*Figure 21* Panel A: Human accuracy: Basic AB (T2 accuracy conditional on T1 correct, T2|T1), raw T1 accuracy, Swaps. Reproduced from Chun and Potter (1995). Panel B: $ST^2$ model accuracy: Basic AB (T2 accuracy conditional on T1 correct), raw T1 accuracy, swaps.

The $ST^2$ model is capable of replicating a range of behavioural data related to the AB (Figure 21). The basic AB describes the accuracy at detecting T2 conditional on T1 being correctly reported, when T1 and T2 are embedded in a continuous RSVP stream of distractors. Raw T1 accuracy describes the accuracy of reporting T1 irrespective of whether T2 was seen or not. If T1 and T2 are correctly identified but reported in the wrong order, this condition is referred to as a swap.

### 3.3.4 Episodic $ST^2$ model

The $ST^2$ model fails to accommodate the recent findings of people not showing an AB if multiple targets are presented in a row (spreading the sparing and whole report; see Section 2.3.2). The need to adapt the previously published $ST^2$ model in order to account for recent findings prompted the development of the episodic $ST^2$ model ($eST^2$ model; Wyble et al., 2009).

Th $eST^2$ model's most pronounced new feature (in comparison to the previously published $ST^2$ model) is a competitive regulation of attention between types, tokens and the

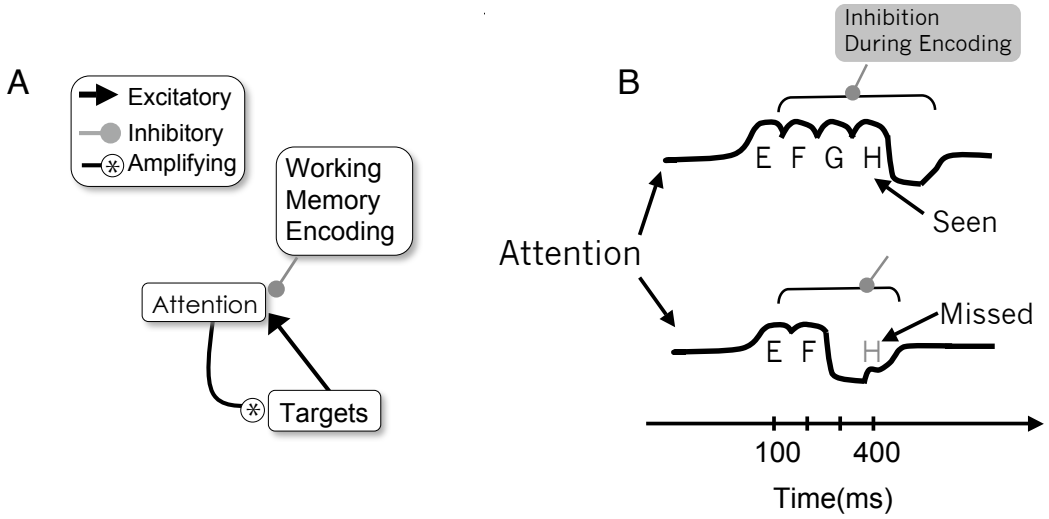


*Figure 22* Panel A: Competitive regulation of attention; bottom-up target input attempts to trigger (excites) attention while the working memory encoding process attempts to shut off (suppresses) attention. Panel B upper figure: Activation dynamics of the eST$^2$ model when multiple targets are presented in a row (i.e. whole report). Panel B lower figure: Activation dynamics of the eST$^2$ model during the AB. Adapted from Wyble et al. (2009).

blaster (named 'attention' in Figure 22A). Figure 22A illustrates how the deployment of attention (through the blaster) to target items is determined by both excitation from bottom-up input and inhibition from working memory encoding.

A second innovation of the eST$^2$ model concerns the nature of tokens. Normally, the ST$^2$ model - as published in Bowman and Wyble (2007) - attempts to assign one type per token. At lag 1, however, two targets can bind to a single token, yielding high target detection accuracy at the cost of a loss of order information (see Section 3.3.3). When applying this implementation to sparing of more than two consecutive targets, this predicts that observers would not be able to recall any order information, which is not the case (Wyble et al., 2009). The eST$^2$ model solves this problem by employing the strict rule of binding one type to one token. If multiple targets arrive before a previous encoding process has completed, the types are encoded into tokens in a staged fashion based on type trace strength. The strongest target is encoded into token 1, the second strongest target is encoded into token 2 and so on. Hence, unlike the previous ST$^2$ model, the eST$^2$ model proposes that, at lag 1 in a 'regular' AB-experiment, T1 and T2 are encoded into separate tokens. However, depending on their trace strength, the type-token assignment might be T1 $\rightarrow$ token1 and

T2 → token2 if T1 is stronger than T2, or T2 → token1 and T1 → token2 if T2 has more trace strength than T1.

#### How the eST$^2$ model blinks and also spreads the sparing

Figure 22B illustrates how the eST$^2$ model explains (a) high accuracy if multiple targets are presented in a row and (b) low accuracy during the AB, if multiple targets are separated by a gap of at least 200ms (and no more than 700ms) in target input.

**(a)** Figure 22B (top) shows the activation dynamics of the eST$^2$ model when multiple targets are presented in a row. As there is continuous bottom-up input from targets, these win the competition to excite attention over the working memory encoding process, which is trying to suppress attention. Consequently, all targets receive attentional enhancement from the blaster and have enough activation to be encoded into working memory, which results in all targets being seen.

**(b)** Figure 22B (bottom) shows the activation dynamics of the eST$^2$ model during the AB. For three targets separated by a temporal gap, an AB occurs (see Section 2.3.2). As no targets are presented during the temporal gap, there is no excitation from bottom-up input to attention (the blaster) and the competition is lost. Due to the inhibition from the working memory encoding process of the previous targets, attention (the blaster) goes below a critical threshold and the final target's bottom-up input cannot sufficiently excite the blaster. The target fails to receive the attentional enhancement from the blaster and cannot be encoded into working memory, which results in this target being missed during the AB.

### 3.3.5 Neural activation patterns from the ST$^2$ model

Neural activation traces extracted from the ST$^2$ model are referred to as virtual ERPs. The methodology for generating virtual ERPs from the ST$^2$ model is described in detail in Chapter 4.

## 3.4 Summary

In this chapter, we have reviewed current theories of the AB to provide the background for the work that is presented in this thesis. The theories of the AB can be separated into three groups based on their similarities. Firstly, theoretical accounts that hold the distractor following T1 responsible for the AB (see Section 3.1.1 and 3.1.4). Second, theories that suggest competition between T1 and T2 throughout the AB, such as the interference or resource sharing theory (see Section 3.1.3). And third, two-stage theories that argue for serial working memory encoding during the AB (see Section 3.1.2).

These theoretical frameworks have inspired a number of formal models of the AB (see Section 3.2), which simulate behavioural data related to the AB. For each of the formal models that are implemented in neural network architectures, we have analysed the extent to which their neural activation traces can be associated with human EEG data. For some models, the activation patterns show a qualitative resemblance to the human EEG traces recorded during the AB. Nevertheless, due to the discrepancies, they mostly seem inadequate as a tool for generating detailed predictions about human EEG (and ERP) data.

In Section 3.3, this chapter concluded with a description of the $ST^2$ model, which is the basis for much of the work presented in the rest of this thesis. In the next chapter, we provide a detailed description of this thesis' methodological contribution, which is to generate virtual ERPs from the $ST^2$ model. We show how our novel approach is capable of a more precise replication of the human EEG/ERP data and, consequently, enables us to make detailed predictions about experimental data, which can be used to validate the theory underlying the $ST^2$ model.

# Chapter 4

# Virtual ERPs from the $ST^2$ model

Chapters 2 and 3 reviewed the literature relevant to the work presented in this thesis. This chapter introduces a novel approach, which is the methodological contribution of this thesis. We explain how artificial 'electrophysiological' traces, so-called virtual ERPs, are extracted from the $ST^2$ model. Computational models commonly replicate behavioural data, so there is no established methodology for generating virtual ERPs. In Section 4.3, we propose a method for this additional dimension to computational modelling. We justify our methodology by using the most straightforward procedure and keeping our approach as close as possible to the mechanisms that are assumed to underlie the human ERP. Section 4.5 describes the notion of virtual ERP components, which - in analogy with human ERP components - are associated with various stages of cognitive processing.

## 4.1 Motivation

Computational modelling of cognition is commonly focused on the replication of behavioural data. In this domain, computational models have proven to be a useful tool for illustrating cognitive theories and construing experimental predictions. However, as discussed in Section 1.1.2, cognitive science is not limited to behavioural experiments. Recent advances in brain imaging techniques along with affordable powerful computers allow researchers to monitor the participant's behaviour at a certain task and, in addition, record the underlying brain activity. EEG, a technique with particularly high temporal resolution, allows the recording of neural activity from the participant's brain while he or she is performing

a given task. An experiment can thus be designed to investigate cognitive processes that the participant might not even be consciously aware of, thus complementing the analysis of behavioural data.

As described in Section 2.2, the goal of computational modelling of cognition is to add a formal framework to textual cognitive theories. An implementation of these formal models using neural network techniques provides a powerful tool for (a) validating the cognitive theory and (b) making experimental predictions from the model. In terms of behavioural results, neural network models are commonly assessed by testing their ability to replicate human behavioural accuracy scores under various conditions. Some computational models also simulate reaction times (e.g. Cohen, Dunbar, & McClelland, 1990), which provide an additional means of validating the model.

However, as psychological research is not restricted to the behavioural side of experimentation - why should this be the case for computational modelling? In this chapter, we propose an additional dimension to computational modelling. We utilise the fact that cognitive neural networks consist of nodes, which are inspired by the biological characteristics of neurons in the brain. A neural network node's activation states can be treated as the analogue of the activation of an assembly of neurons in the human brain. When monitoring neural network node activation over time, these traces can be compared to corresponding neural activation patterns from the human brain, as expressed by the EEG. Akin to the technique of generating event-related potentials (ERPs) from continuous human EEG (see Section 2.1), we extract model activity related to the onset of an event. Hence, the traces that are generated from model activity are referred to as virtual ERPs (or vERPs), whereas the ERP extracted from raw EEG recorded at the human scalp is henceforth referred to as human ERP (or hERP).

## 4.2 Hypothesis

Figure 23 summarises the hypothesis this work aims to test. As depicted in the illustration, the literature holds a set of behavioural data for the particular experimental observation (i.e. the AB). Based on these data, Bowman and Wyble (2007) proposed a computational model, which describes a cognitive theory and was designed to simulate human behavioural

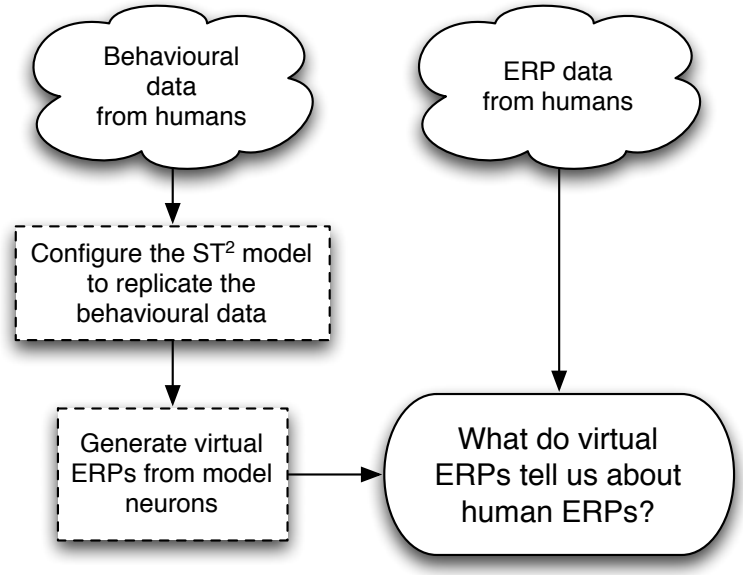


*Figure 23* The hypothesis underlying the virtual ERP approach.

accuracy scores in the AB task using its neural network implementation. We can also gather electrophysiological (hERP) data related to the given experimental phenomenon. The question is thus: If we generate activation traces from nodes of the neural network model, can we produce meaningful vERP traces that allow us to replicate and make predictions about hERPs in the way traditionally done with behavioural data?

This project aims to test this hypothesis with respect to the AB using the $ST^2$ model (Bowman & Wyble, 2007). As shown in Section 2.3.2, there is a large amount of experimental literature investigating the AB both using behavioural and EEG/ERP methods. As depicted in Figure 21 and discussed in Section 3.3.3, the $ST^2$ model replicates a wide range of behavioural results from the AB literature. In line with our hypothesis from Figure 23, we want to test what the vERP traces generated from the $ST^2$ model's neural network node activation traces can tell us about hERP patterns observed during the AB. We will assess the vERP traces from the $ST^2$ model in terms of how well they replicate the time course and profile of the hERP data. This allows us to validate the $ST^2$ model. Furthermore, we will explore the possibilities of using the vERP technique to make new predictions about experimental results that could be verified in future experiments.

## 4.3 A methodology for virtual ERPs

Due to the novelty of this approach, there is no established methodology for generating vERPs. In the context of the AB, the neural network models discussed in Section 3.2.2 (Fragopanagos et al., 2005) and 3.2.3 (Nieuwenhuis, Gilzenrat, et al., 2005) generate neural activation patterns from the membrane potentials of specific neural network nodes. For the spiking neuron approach (see Section 3.2.1; Dehaene et al., 2003), the firing rate of neural network nodes is taken as a measure of neural activation.

### 4.3.1 Node potentials in the ST$^2$ model

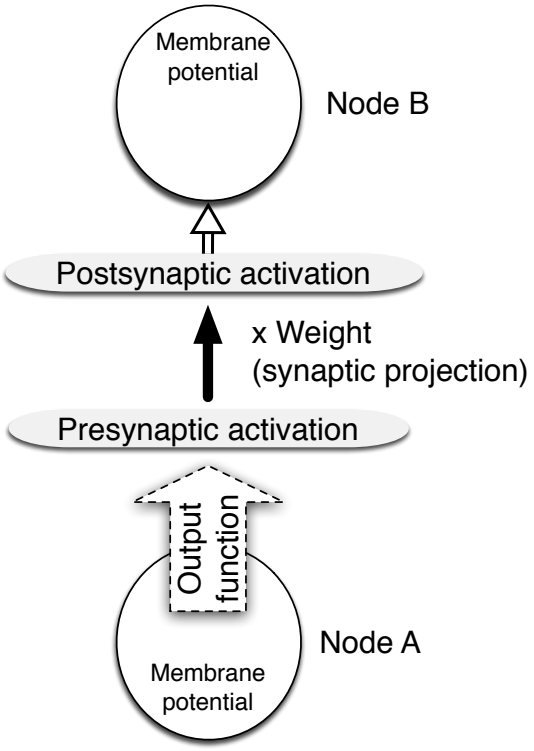


*Figure 24* A typical pair of neural network nodes situated in two neighbouring connected layers of the ST$^2$ model. As shown in the figure, we can extract the membrane potential, presynaptic activation and postsynaptic activation for each node of the ST$^2$ model's neural network implementation (neural-ST$^2$).

As illustrated by the model architecture in Figure 20, the ST$^2$ model consists of a number of layers, each containing several nodes. Nodes of one layer are connected to nodes in other layers via excitatory and inhibitory connections. In order to describe the activation

dynamics at the individual node level, Figure 24 focusses on a single pair of nodes.

The two nodes depicted in Figure 24 represent a typical $ST^2$ node pair in two connected neighbouring layers. Nodes in the $ST^2$ model receive input from other layers via weighted connections and update their membrane potential according to shunting equations based on the Hodgkin-Huxley approach (Hodgkin & Huxley, 1952). Once the membrane potential reaches a given threshold, they produce output according to a rate-coding X-over-X+1 function (O'Reilly & Munakata, 2000).

Figure 24 summarises the activation dynamics of a typical node in the $ST^2$ model as follows. The membrane potential describes the activation within the node. The weighted connections between nodes are assumed to correspond to major synaptic projections in the brain. We refer to activation output by a node as presynaptic activation and after multiplying the presynaptic activation by its corresponding weight, the result is referred to as postsynaptic activation.

### 4.3.2 Neural correlates of human ERPs

The difference in electric charge between the dendrite and the postsynaptic cell body of an active neuron creates an electric dipole. To generate a signal that is strong enough to be registered by the EEG, a population of neurons has to be active together and spatially aligned, which causes the individual dipoles to summate. Cortical pyramidal neurons have long-range connections and are aligned perpendicular to the cortex, which is why these neurons are assumed to be a major contributor of the human EEG (Baillet, Mosher, & Leahy, 2001; Luck, 2005). Pyramidal neurons release glutamate as their neurotransmitter and are therefore primarily excitatory.

### 4.3.3 Choosing a neural network node potential

In generating vERPs, we aim to keep our approach as close as possible to the mechanisms that are assumed to occur in the brain. Hence, unlike the neural network models described in Section 3.2, we do not use membrane potentials. Instead, we choose to generate vERPs by summing over postsynaptic activation values (as depicted in Figure 24), the rationale being that it is postsynaptic potentials between the dendrite and the cell body that generate the

EEG signal (as discussed in Section 4.3.2). In line with pyramidal neurons forming mainly excitatory connections in the brain, the vERP consists of postsynaptic activation values from excitatory connections only. Note that only activation traces from connections between layers (and not self-loops that connect nodes within a layer) are included in the vERP, as these are assumed to be an analogue of long-range pyramidal neurons that contribute towards the signal measured in the hERP.

### 4.3.4 Virtual ERP averaging procedure

We adopt the most straightforward approach and sum over all nodes of a given subset of layers in order to avoid a specific weighting of layers or normalisation setting, which would be difficult to quantify from a neurobiological perspective.

In addition, neurophysiological evidence suggests that there is a delay of around 70ms for neural activation related to visual processing to travel from the retina to occipital areas (Schmolesky et al., 1998). To account for this delay, vERPs are shifted by 70ms.

### 4.3.5 Virtual grand average ERPs and virtual ERPimages

The shortcomings of the grand average ERP technique (as discussed in Section 2.1.1) also apply when generating virtual ERPs. A simulation of the $ST^2$ model contains a number of trials encompassing a range of target strength values. A grand average vERP will illustrate the average time course of activation between conditions but is blind to inter-trial fluctuations. As with human ERPs, we can make use of the ERPimage technique and generate virtual ERPimages, which depict the dynamics of the model for each individual trial (i.e. a particular target strength combination) of the simulation.

### 4.3.6 A word of caution

It is obvious, however, that vERPs remain a coarse approximation of the hERP. Some factors, such as the distortion of the signal by the scalp, are not addressed[1]. Due to these

[1]For an example of a biologically informed approach to modelling ERPs that is based on a neurobiologically constrained source reconstruction scheme, see the dynamic causal modelling technique (David, Harrison, & Friston, 2005; David, Kilner, & Friston, 2006).

limitations, one can realistically only expect to obtain a qualitative rather than a quantitative match to the data. Nevertheless, we hypothesise that vERPs from the $ST^2$ model allow us to make sensible predictions about the ERPs recorded from human participants in EEG experiments.

## 4.4 Changes to the $ST^2$ model in comparison to Bowman and Wyble (2007)

For this work we generate vERPs from the $ST^2$ model with as few parameter changes as possible compared to the previously published version of the $ST^2$ model. Table 1 contains a list of the neural network weight values that were modified in the process of this work. Note that we can still reproduce all behavioural data published in Bowman and Wyble (2007).

| Layer1 ⇒ Layer2 | Weight value |
|---|---|
| 100ms SOA: Input layer ⇒ Masking layer | 0.023 (0.022) |
| 50ms SOA: Input layer ⇒ Masking layer | 0.058 (0.05) |
| TFL ⇒ Blaster | 0.02003 (0.018) |
| Blaster recurrent excitation | 0.0112 (0.01) |

*Table 1* List of weights that were modified during this work. The values from Bowman & Wyble, 2007 are shown in brackets. All other parameters remained unchanged.

The number of distractor nodes in stage one is increased from 10 to 15 nodes. This has no effect on behavioural accuracy, but is required to generate 50ms SOA vERP traces as otherwise, due to the fast presentation rate, nodes are not able to decay to baseline before being reactivated.

## 4.5 Generating virtual ERP components

As summarised in Section 2.1.1, the human EEG is often analysed by means of comparing ERP components between conditions. A human ERP component is typically recorded from a particular set of electrodes and associated with certain cognitive processes. In the $ST^2$ model, different layers of the model were designed to correspond to various stages of cognitive processing in the brain. By summing over neural network activation from nodes

within specific layers of the $ST^2$ model, we can extract vERP activity related to particular stages of cognitive processing. In analogy with human ERP components, we can compare how the resulting vERP components are modulated by the various experimental conditions.

### 4.5.1 Early visual processing

Correlates of early visual processing in the hERP are observed at occipital electrode sites, which are located above early visual cortical areas. In the $ST^2$ model, targets and distractors in the RSVP stream are 'presented' to the input layer. The input layer thus corresponds to very early stages of processing in the brain. At the masking layer, each item is subject to competition from neighbouring items, which simulates forward and backward masking at early visual stages. The amount of masking is determined by the strength of the neighbouring items, where an item's strength is the model's analogue of the perceptual features of such a stimulus. Since the next layer above the masking layer - the item layer - is involved with semantic processing, the item layer is conceptually distinct from early visual processing. ERP early components reflect early perceptual processing as they are modulated by changes in visual features of stimuli but they are not effected by higher level manipulations, such as semantic congruency. Consequently, the item layer should not contribute towards a virtual ERP early component. The input and masking layer, however, reflect perceptual processing of stimuli and, thus, most closely resemble processes occurring in early visual cortex.

Figure 25A shows an example vERP trace when summing across input and masking layer. We term this vERP component virtual SSVEP (vSSVEP), as the SSVEP wave is the correlate of early visual processing for repeatedly presented items (for instance during RSVP) in the human ERP (see Section 2.1.1). Since the vERP trace is plotted for the whole RSVP stream, it also reflects the on- and offset of the RSVP stream.

### 4.5.2 Working memory encoding

The P3 component of the hERP is most prominent at parietal electrode sites, however, depending on the experimental paradigm, the P3 can often also be recorded throughout

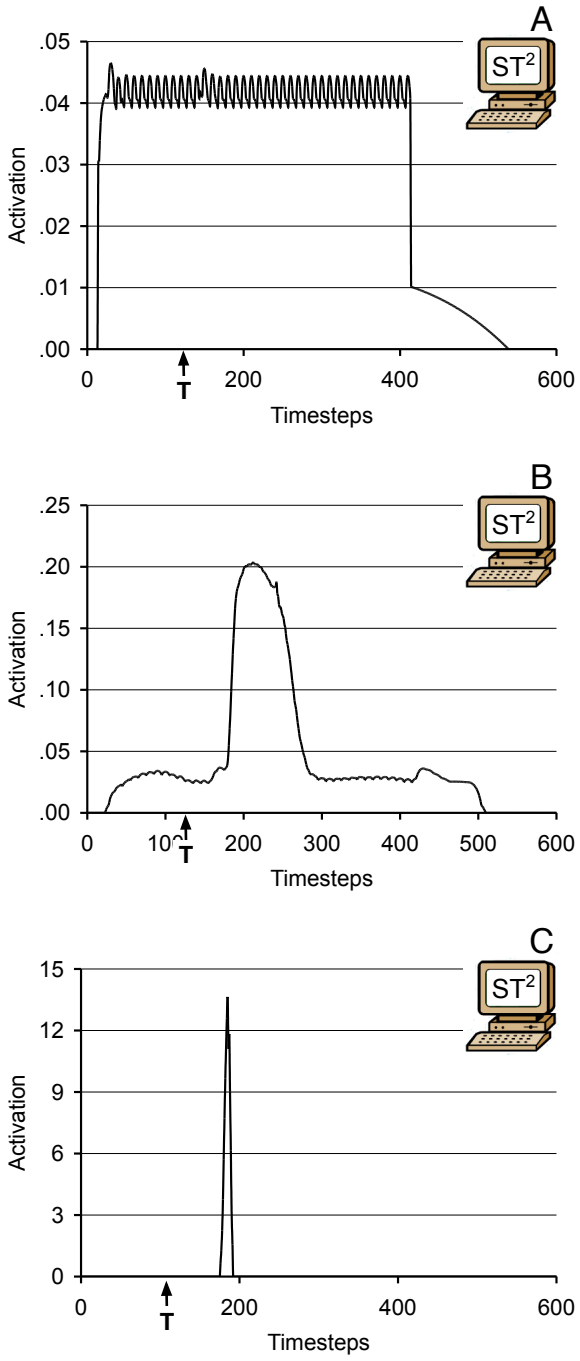


*Figure 25* Panel A: Virtual SSVEP component from input and masking layers resembling early visual processing. Panel B: Virtual P3 component from item layer, TFL, binding pool gate nodes and token gate nodes, resembling working memory encoding. Panel C: Virtual N2pc component from the blaster node reflecting the firing of attentional enhancement, which initiates attentional selection. In all figures, time point 0 corresponds to the onset of the RSVP stream.

occipital, central and frontal electrode sites. Rather than being associated with neural activity from a specific brain region (like the ERP early components for instance), the P3 seems to reflect a 'global brain event' that involves multiple cortical areas from various parts of the brain. Accordingly, we generate the virtual P3 component (vP3) from multiple layers of the $ST^2$ model.

The P3 is considered to be a correlate of working memory consolidation (see Section 2.1.1) and has been associated with conscious perception (see Chapter 8 and also Sergent et al. (2005)). In the $ST^2$ model, working memory encoding occurs by creating a binding link between types from stage one and tokens from stage two, which we refer to as *tokenisation*. The tokenisation process is facilitated by an attentional enhancement from the blaster, which projects to the item layer and the TFL in stage one. Through the blaster, activation in the item layer and the TFL is increased until the activation from the TFL triggers the nodes in the binding pool, which in turn are connected to the token nodes in stage two. Hence, item layer and TFL nodes, nodes in the binding pool and token nodes in stage two are involved in encoding (or tokenising) an item into working memory and, consequently, nodes from these layers contribute towards the virtual P3 component. Figure 25B shows an example virtual P3 component containing activation from later parts of the first stage (item layer and TFL), the nodes in stage two (tokens) and the binding link connecting the two stages.

### 4.5.3 Attentional selection

The N2pc component of the ERP has been associated with attentional selection (see Section 2.1.1). In the $ST^2$ model, attention is modelled through the blaster, which is triggered if items match the target template and provides targets with an enhancement during working memory encoding. In order to generate a virtual ERP component that reflects the firing of the attentional enhancement and the initiation of attentional selection, we sum across activation from the output nodes of the blaster to generate the virtual N2pc component (vN2pc, example vERP trace in Figure 25C). Unlike the vERP components described in the previous sections, the virtual N2pc component does not contain activation from multiple layers but reflects the output from a single neural circuit, i.e. the blaster.

## 4.6 Summary

The current chapter describes the virtual ERP technique that is the methodological contribution of this thesis. We have discussed the methodology that is employed for generating virtual ERPs. The next part presents the experimental results from the EEG studies conducted for this thesis and shows how the vERP can be used to validate the theory underlying the $ST^2$ model. Furthermore, we demonstrate how virtual ERPs can be used to explain the experimental results and make predictions about the results of further experiments.

# Part III

# Using virtual and human ERPs to explore the limits of conscious perception

# Chapter 5

# How distractors influence target selection in RSVP

This chapter investigates the role of distractors on target processing in RSVP paradigms. In a first manipulation, we remove all distractors except the one following the target from the RSVP stream, thus employing the so-called skeletal paradigm. We present EEG results that show how the ERP correlates of early visual processing, target selection and working memory encoding are modulated if a target is presented in skeletal presentation compared to when it is embedded in an RSVP stream of distractors. Subsequently, we modify the $ST^2$ model's architecture to simulate behavioural and virtual ERP data for skeletal presentation. The simulations allow us to hypothesise about the differences in target processing between RSVP and skeletal presentation. Second, we remove only the distractor following the target in the RSVP stream. This manipulation affects the ERP correlates of working memory encoding. The $ST^2$ model replicates the human data in terms of simulated accuracy and virtual ERPs and suggests that the experimental effects are due to differences in bottom-up target strength. Finally, we discuss how the exploration of target processing in RSVP is important for assessing theories of the AB.

## 5.1 Introduction

In this chapter, we investigate how distractors in the RSVP stream influence the strategy that observers employ to select targets and encode them into working memory. To this end, we manipulate the context of target presentation by selectively removing distractors from the RSVP stream and analyse how this influences target processing.

### 5.1.1 RSVP without distractors: Skeletal presentation

In the first part of the chapter, we compare the EEG signatures of visual processing for target items followed by just a single mask - using the *skeletal RSVP* paradigm (Ward, Duncan, & Shapiro, 1996) - to targets in regular RSVP streams. The skeletal RSVP[1] task is a paradigm in which the presentation stream contains only targets and their following backward masks. Hence, a characteristic trait of regular RSVP, i.e. the continuous stream of distractors surrounding the target, is not present in the skeletal task. See Figure 26 for an example of how the skeletal presentation paradigm differs from a regular RSVP stream.

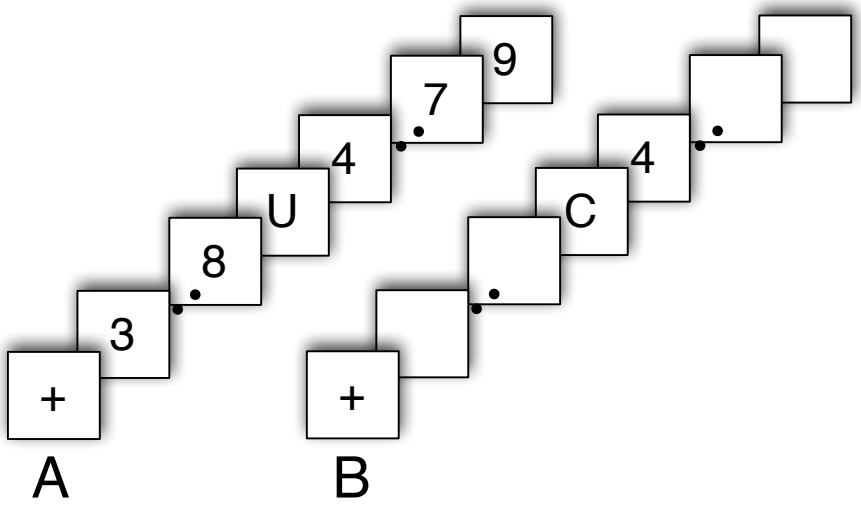


*Figure 26* Panel A: A regular RSVP stream where a target letter is embedded in a stream of digit distractors. Panel B: The skeletal presentation paradigm, which contains only the target letter and the following digit distractor as its mask.

The skeletal paradigm that is often used to study the AB was originally derived from a

[1] referred to as *skeletal* from now on

spatial paradigm used in Duncan, Ward, and Shapiro (1994). In their study, Duncan et al. (1994) presented two targets, which were both followed by pattern masks, at different spatial locations and it was found that it took participants several hundred milliseconds to switch from one target to the other. This was far more than what conventional theories of spatial attention predicted and the delay was hypothesised to be due to masking effects (Ward et al., 1996; Moore, Egeth, Berglan, & Luck, 1996). In a later study, Ward, Duncan, and Shapiro (1997) noted the similarity in the time course of interference between masked targets *in space* (in the Duncan et al. (1994) study) and masked targets *in time* as exemplified by the AB in regular RSVP paradigms. Ward et al. (1997) speculated that both effects might be due to the same underlying mechanism, i.e. a cognitive system with limited capacity. Accordingly, in both Duncan et al. (1994)'s spatial paradigm and during the AB in RSVP, the second target arrives while the visual system is processing the first target and thus accuracy at detecting the second target is reduced. However, the spatial paradigm from Duncan et al. (1994) requires not only a switch from one target to the other but also necessitates a redistribution of attention to a different location in space. Hence, in order to make the spatial paradigm more similar to the (non-spatial) AB task, Ward et al. (1997) employed the skeletal task, where although all items appear in the same spatial location, the presentation stream contains only the targets and the distractors (masks) that follow each of the targets.

### The attentional blink in skeletal presentation

Masking functions show how the occurrence and duration of a target's mask influence the accuracy at detecting that target (see e.g. Breitmeyer, Ro, & Ogmen, 2004). Accordingly, it is generally found that forward masking has a smaller detrimental effect on target accuracy than backward masking (Enns & Di Lollo, 2000). In line with this, Breitmeyer et al. (1999) show that the removal of the distractor immediately preceding T1 has no effect on T2 accuracy during the AB, which suggests that T1's forward mask has only a small influence on the AB.

So does the skeletal paradigm produce an AB effect comparable to that observed when employing regular RSVP paradigms? In skeletal presentation, the targets are backward masked but there is no forward masking. As seen in Figure 27, skeletal presentation

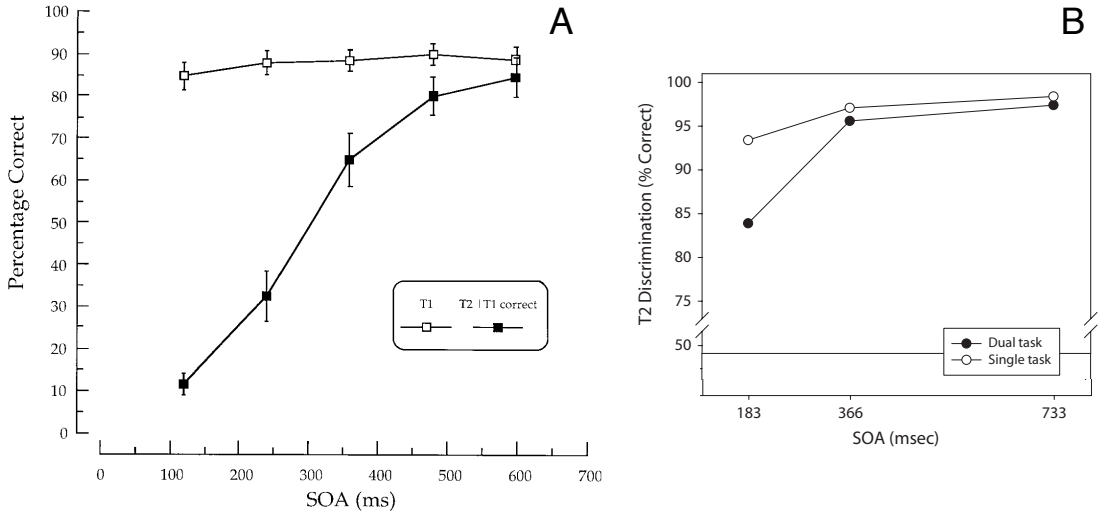


*Figure 27* Behavioural accuracy scores from AB studies using the skeletal paradigm. Panel A: White circles (T1) show T1 accuracy per lag. Black squares (T2|T1 correct) indicate T2 accuracy per lag conditional on T1 being correct. Adapted from McLaughlin et al. (2001). Panel B: White circles (Single task) show T2 accuracy per lag when subjects were instructed to ignore T1 and report T2. Black circles (Dual task) show T2 accuracy per lag when subjects were instructed to report both T1 and T2 per lag. Adapted from Rolke et al. (2007).

produces an impairment in the detection of T2, which is similar to the AB in regular RSVP (McLaughlin et al., 2001; Rolke et al., 2007). However, the AB curve in skeletal presentation is considerably shorter in duration. As seen in Figure 27, the AB curve for skeletal presentation abides for 300-400ms after T1 presentation. In regular RSVP, however, the AB typically lasts for approx. 600ms post T1 presentation. Furthermore, the skeletal AB does not show lag 1 sparing, instead, the lowest T2 performance during the AB in skeletal presentation is, in fact, at lag 1 (see Figure 27).

### Distractor effects on target processing strategies

Figure 27 shows that, although there are some discrepancies, both skeletal presentation and regular RSVP paradigms seem to produce an AB effect. Nevertheless, we argue that whether or not distractors surround the target in the presentation stream influences how this target is processed. Consequently, there should be differences in the EEG correlates of target processing between skeletal presentation and regular RSVP.

Targets in RSVP paradigms are embedded into a continuous stream of distractors and

it requires a category-level task filter mechanism to select targets from distractors. In skeletal presentation, however, targets are not preceded by distractor items and are thus marked by visual onset. The visual system is likely to process continuous streams differently from visual onsets. Hence, although the behavioural AB data show a similar pattern, the underlying mechanisms of how targets are processed might be different.

### 5.1.2 Reducing backward masking in RSVP

In the second part of the chapter, we investigate the effect of removing the distractor following the target in the RSVP stream. A mask following a target impairs the accuracy at detecting that target and the strength of the impairment is determined by the type of mask employed (Enns & Di Lollo, 2000). Although certain pattern masks are particularly strong, masking effects also occur if a target letter is followed by a digit distractor. With respect to the AB, various experiments have demonstrated that the AB is strongly modulated by the amount of backward masking from the distractor following each of the targets (Chun & Potter, 1995; Giesbrecht & Di Lollo, 1998; Vogel & Luck, 2002).

The second experimental manipulation presented in this chapter is essentially an inverse of the first manipulation. We remove only the distractor following the target in the RSVP stream while keeping everything else identical to a regular RSVP paradigm and refer to this type of paradigm as a T+1 blank stream. Through this manipulation, we can analyse the effect of reduced backward masking on target processing in RSVP.

### 5.1.3 Motivation and overview

The investigation of the AB requires a detailed knowledge of how targets are processed in RSVP. This chapter investigates the role of distractors for targets in RSVP using electrophysiological methods and compares the results to simulations from the $ST^2$ model. As our paradigm contains only one target, we take a step back from two target paradigms investigating the AB. In such a single target paradigm, there is no interference between targets, which reduces potential confounds when analysing target processing in skeletal, RSVP and T+1 blank streams.

## 5.2 Methods

This chapter analyses behavioural and EEG data from Experiment 1, which contains the regular RSVP, skeletal and RSVP T+1 blank conditions. Please refer to the appendix for a detailed overview of the methods employed in Experiment 1. The methods specific to the EEG analyses presented in this chapter are described in the following section.

### 5.2.1 EEG methods

In this chapter, we analyse EEG data from occipital-parietal scalp locations, more precisely, the P7 and P8 electrode sites. We average across these two sites as we are not interested in lateral effects but focus on ERP components that are not specific to one of the hemispheres. Unlike the other chapters, where we analyse data from the Pz electrode site when investigating the P3 component, the analysis here also focuses on early visual processing. The ERP trace averaged across the P7 and P8 electrodes contains both the P3 component and ERP components associated with early visual processing, which is why these electrode sites were chosen for this chapter's analyses.

All ERPs contain only those trials in which the target was correctly identified. After artifact rejection (the details are described in the appendix), the skeletal condition contains 554 trials, the RSVP condition contains 1,574 trials and the T+1 blank condition contains 1,819 trials.

## 5.3 Target processing in skeletal presentation

We first present the behavioural and electrophysiological results for targets in skeletal presentation and targets in RSVP. Following this, we elaborate on how the $ST^2$ model can be modified in order to simulate skeletal presentation and conclude this section with a theoretical discussion.

### 5.3.1 Results: Behaviour

Overall, when compared to RSVP, skeletal presentation makes targets easier to detect. Participants report 72% (SEM 4) of targets correctly if they are embedded in a regular RSVP

stream, whereas in skeletal presentation target accuracy is 81% (SEM 4). The difference in accuracy scores between targets in RSVP and skeletal presentation is significant; $F(1,19) = 10.7$, MSE $< 0.01$, $p = 0.004$.

### 5.3.2 Results: Electrophysiology

#### ERP early components

Whether a target is presented in skeletal presentation or RSVP has a strong effect on early processing. Figure 28A illustrates a highly significant difference in the P1 and N1 ERP early components between targets in RSVP and skeletal presentation. The mean absolute value in the area from 0-200ms after target presentation is 1.0$\mu$V (SEM 0.1) for RSVP targets and 3.5$\mu$V (SEM 0.3) for targets in skeletal streams ($F(1,19) = 103.1$, MSE = 0.5, $p < 0.001$).

Instead of evoking P1/N1 early components, RSVP targets produce an ssVEP wave oscillating at the same frequency as the presentation rate of items in the RSVP stream (Figure 28B). As each item is presented for 47.1ms (corresponding to the RSVP rate of roughly 20 items per second), this results in a peak at approx. 21Hz in the FFT plot.

#### ERP P2 component

As seen in Figure 28A, the ERP for skeletal targets shows a positive P2 wave between 200 and 300ms, which is followed by the P3 component. The ERP for RSVP targets, however, does not deviate from baseline until the onset of the P3 component. The difference between skeletal and RSVP targets in the mean value of the 200-300ms window is significant ($F(1,19) = 20.6$, MSE = 8.7, $p < 0.001$). For skeletal targets, the P2 has an average size of 3.6$\mu$V (SEM 1.1). For RSVP targets, the P2 is negligible (mean value -0.6$\mu$V, SEM 0.4).

#### ERP P3 component

The P3 component, which is depicted in Figure 28A, shows a different profile for skeletal compared to RSVP targets. The 50% area latency of the 300-1050ms window is shorter for skeletal (mean 427ms, SEM 21) than RSVP targets (mean 500ms, SEM 19). This difference is significant; $F(1,19) = 10.5$, MSE = 5006, $p = 0.004$. Furthermore, the mean value in

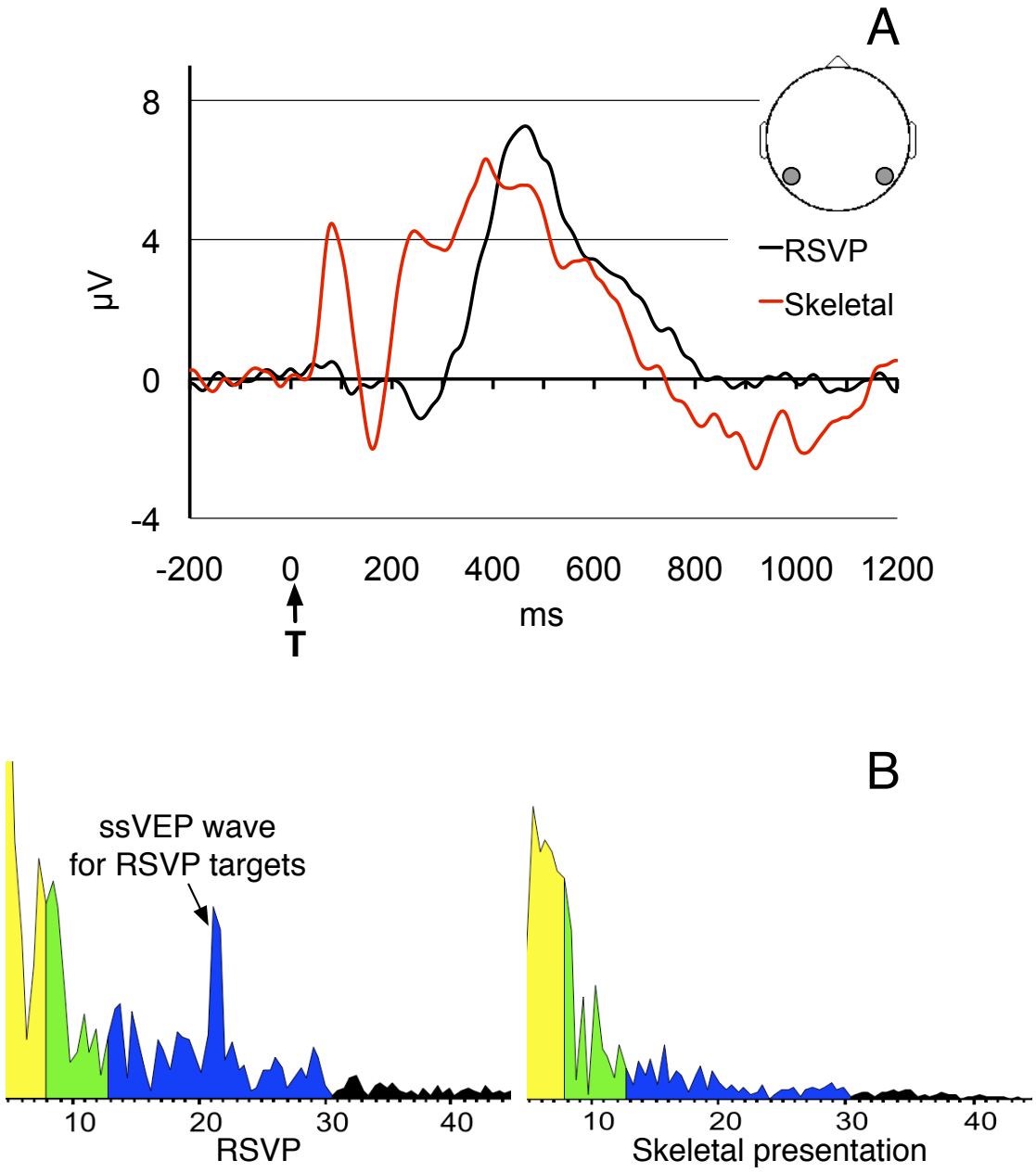


*Figure 28* Panel A: Human ERP for targets in RSVP and skeletal presentation averaged across P7 and P8 electrode locations. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target. Positive is plotted upwards. Panel B: Fast fourier transform (FFT) of the ERP from the P7 and P8 electrode locations for the RSVP (left) and skeletal (right) condition. The RSVP condition shows a peak in the FFT plot at the frequency of target presentation (approx. 21Hz), which is not present for skeletal presentation.

the P3 window for skeletal targets (1.5$\mu$V, SEM 0.4) is smaller than for targets in RSVP (2.3$\mu$V, SEM 0.4); F(1,19) = 4.5, MSE = 1.3, p = 0.047.

### 5.3.3 Modelling skeletal presentation

The ST$^2$ model as published in Bowman and Wyble (2007) and described in Section 3.3 cannot simulate skeletal presentation. In the following, we will show how, by making a number of theoretically justified changes to the architecture of the model, we can replicate our experimental results with respect to skeletal presentation in both the behavioural and the EEG domain.

#### Step 1: Setting distractor values to zero

**Manipulation** In skeletal presentation, the presentation stream contains just the target and the distractor following the target. All other distractors are removed and replaced by blank intervals. In order to simulate such a stream in the ST$^2$ model, we modify the array of values that serves as input to the model. All distractors - except the one following the target - are set to a value of zero, which corresponds to no activation.

**Results** The modification of the input array has a strong effect on virtual ERP traces resembling early visual processing (see Section 4.5.1 for a corresponding description of vERP methodology). For targets in RSVP, the model shows a continuous *virtual ssVEP* wave oscillating at the frequency of target presentation (Figure 29), hence, the model replicates the human data from Figure 28. The first item of the RSVP stream causes an increase of activation in early layers of the model. As the following items appear in rapid succession, activation in these layers does not decay back to baseline. Rather, the inhibition between items in the masking layer causes layer activation to oscillate around a certain value until the end of the RSVP stream.

In skeletal presentation, there are no distractors and hence there is no activation preceding the target. When the target is presented, this creates a strong burst of activation at early layers of the ST$^2$ model. As there is no forward masking (i.e. the target representation is not inhibited by distractors preceding the target), the activation evoked by the skeletal target at early layers is higher than in regular RSVP. The distractor following the target in

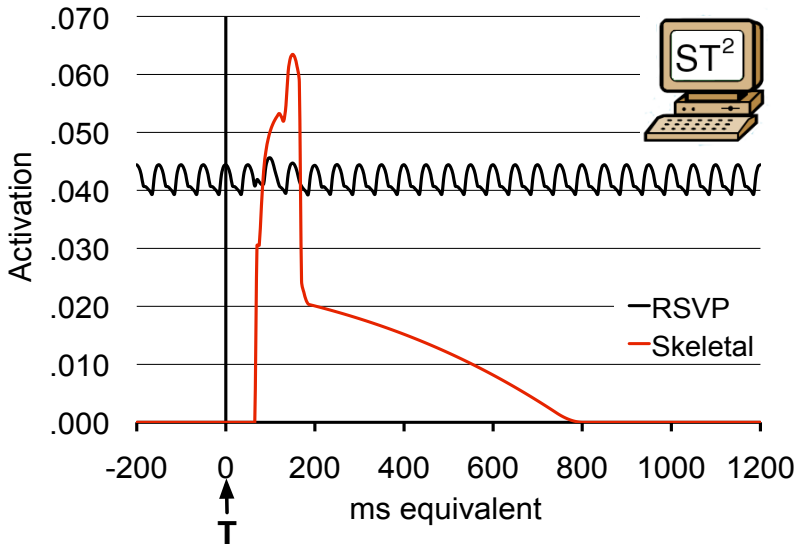


*Figure 29* A virtual ssVEP wave for the RSVP and vERP early components for the skeletal condition from input and masking layer of the $ST^2$ model. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target.

skeletal presentation then produces a second large burst of activation, as its activation at early layers is not constrained by backward masking. All of this activation at early layers occurs between the model equivalent of 100 and 200ms following target presentation. There is thus a qualitative match between the virtual ERP from the $ST^2$ model (Figure 29) and the human ERP early components (P1/N1 wave) for skeletal presentation from Figure 28.

To summarise, virtual ERP activation associated with early visual processing shows a distinct activation for skeletal targets and an oscillatory pattern for RSVP targets, thus qualitatively replicating the human ERP. Furthermore, the timing of the skeletal vERP activation occurs within a similar time window as the P1/N1 wave observed for skeletal targets in the human ERP.

### Step 2: Moving the blaster 'trigger' to masking layer

**Manipulation** The adjustment of the input array for skeletal presentation has the desired effect on virtual ERP activation resembling early visual processing. A replication of behavioural accuracy and the virtual P3 component, however, requires theoretically justified

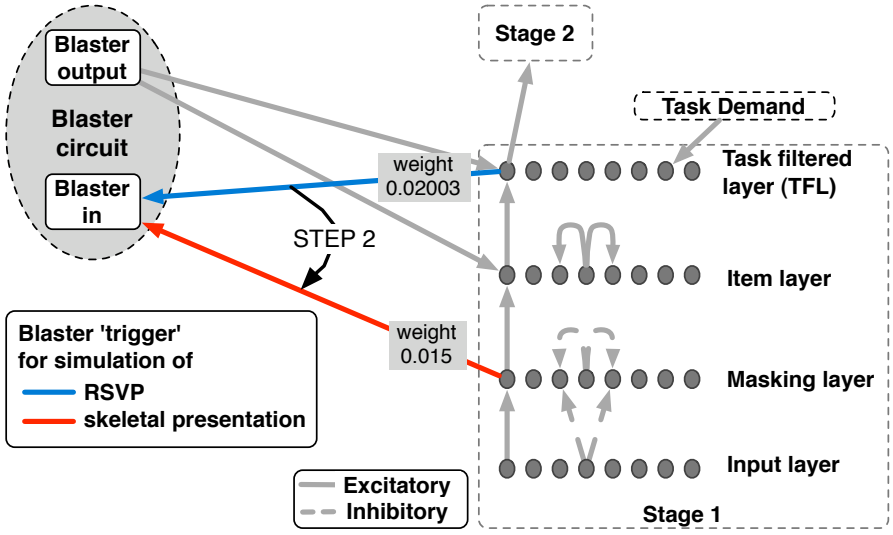


*Figure 30* Step 2 of simulating skeletal presentation. As indicated in the figure, the connection from stage one that triggers the blaster is moved from the task filtered layer to the masking layer.

changes to the architecture of the ST$^2$ model.

As discussed in the introduction, skeletal targets appear as a visual onset on a previously blank screen, whereas in RSVP, the target has to be selected from a continuous stream of distractors. In terms of the ST$^2$ model, we hypothesise that this influences the way in which the blaster is triggered:

- In RSVP, the system cannot distinguish targets from distractors until they have reached the task filtered layer (TFL). In the TFL, the task demand mechanism selects targets by means of selective excitation to target nodes and inhibition to distractor nodes.
- In skeletal presentation, there are no distractors preceding the target, hence, the system can assume that the first item that is 'presented' to the input layer is the target. Accordingly and as seen in Figure 30, we propose that in skeletal presentation the blaster is triggered as soon as activation reaches the masking layer[2]. Moving the connection to the

[2]For the purpose of simulating the skeletal paradigm (i.e. with no distractors preceding the target), our manipulation produces the desired effect. However, our modification of the model architecture would have to be reconsidered in order to simulate a slightly different stream setup, for instance, if the target was also preceded by distractor items (e.g. a stream of the type 'D D T D'). Under these circumstances, the distractors can potentially also fire the blaster, as task demand does not operate until the TFL and, hence, the system cannot distinguish targets from distractors at the masking layer. Note, however, that although distractors can fire the blaster in skeletal presentation, task demand at the TFL will prevent distractors from being tokenised.

blaster from TFL to the masking layer also requires a modification of the weight value of that connection (see Figure 30). This is a technical requirement that is necessary because of the way the ST$^2$ model is implemented[3].

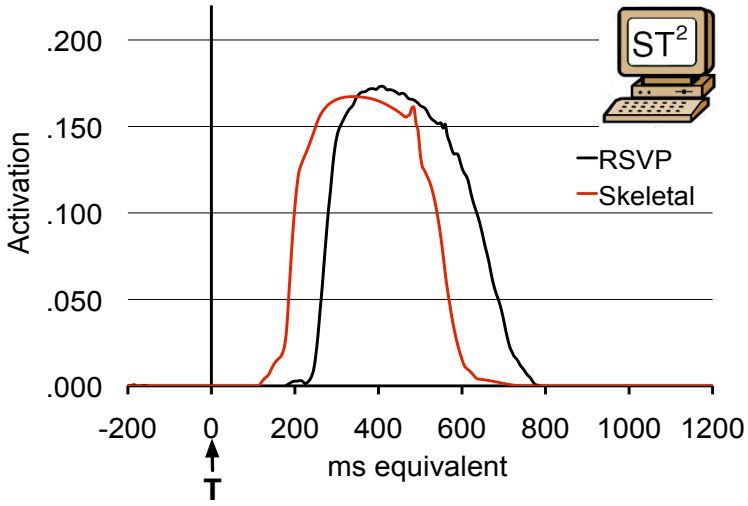


*Figure 31* After Step 2: Virtual P3 for the RSVP and skeletal condition. The RSVP vERP is baseline corrected to -200 to 0ms with respect to target onset to account for distractor related activity, which is absent in the skeletal RSVP condition. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target.

**Results** As activation propagates through the ST$^2$ model, there is temporal delay. Hence, if the blaster is triggered from the masking layer, the blaster fires at an earlier time point than if activation has to propagate to the TFL before the blaster can be triggered. Consequently, the blaster's output to item layer and TFL also occurs earlier in time.

The first consequence of this change is a shift in latency of the virtual P3 (see Section 4.5.2 for a corresponding description of vERP methodology) for skeletal compared to RSVP targets, as seen in Figure 31. In RSVP, the target reaches the TFL and triggers the blaster once the target has been identified as such by the task demand mechanism. Once the blaster becomes active, it can enhance the target for tokenisation. With the change

[3]Compared to the TFL, activation values at the masking layer are higher in absolute terms. Hence, we reduce the weight values between masking layer and blaster to prevent the blaster circuit from being overcharged by the input from the masking layer.

in model architecture for skeletal presentation, the blaster is triggered earlier, thus, it is active and ready when targets reach the TFL. The consequence is an earlier tokenisation (and virtual P3) for skeletal targets.

The change in model architecture means that the blaster will now fire for all skeletal targets, regardless of their input strength. This increases the accuracy of the $ST^2$ model at detecting skeletal targets. Whereas RSVP targets have an average simulated accuracy of 77%, the earlier blaster response caused by the modification of the $ST^2$ model's architecture produces a simulated skeletal accuracy of 100%. Although skeletal behavioural accuracy should indeed be above RSVP accuracy, this is not a good replication of the human behavioural performance for detecting skeletal targets, which is below ceiling. Furthermore, the virtual P3 lacks the distinctive difference in size between skeletal and RSVP targets that is evident in the human P3 component.

A further modification of the weight value between masking layer and blaster does not have the desired effect on simulated accuracy and virtual P3 for skeletal targets. This is due to the blaster 'trigger' functioning in an 'all-or-none' fashion, hence, the weight value would have to be reduced to close to zero before there is any further effect on the target's tokenisation process. Reducing the weight to close to zero, however, has a counterproductive effect as, in this case, the blaster can only be triggered by those targets with the highest strength values. Consequently, only a few targets are tokenised and all other targets are not 'detected' by the model. This reduces the simulated skeletal accuracy to below that for RSVP targets, which is obviously not a desirable replication of the human data either. Consequently, we need to perform one additional modification to the architecture of the $ST^2$ model (as described in the following section) in order to accurately simulate skeletal presentation.

#### Step 3: Decreasing the strength of the task demand mechanism

**Manipulation** An RSVP stream consists of one or more targets embedded in a stream of distractors. In skeletal, however, the stream contains only the target and the following distractor. When an RSVP target arrives at the TFL, the task demand mechanism plays a vital role in selecting the target from simultaneously active distractors. In skeletal presentation, however, the target competes with only one other distractor at the TFL and hence

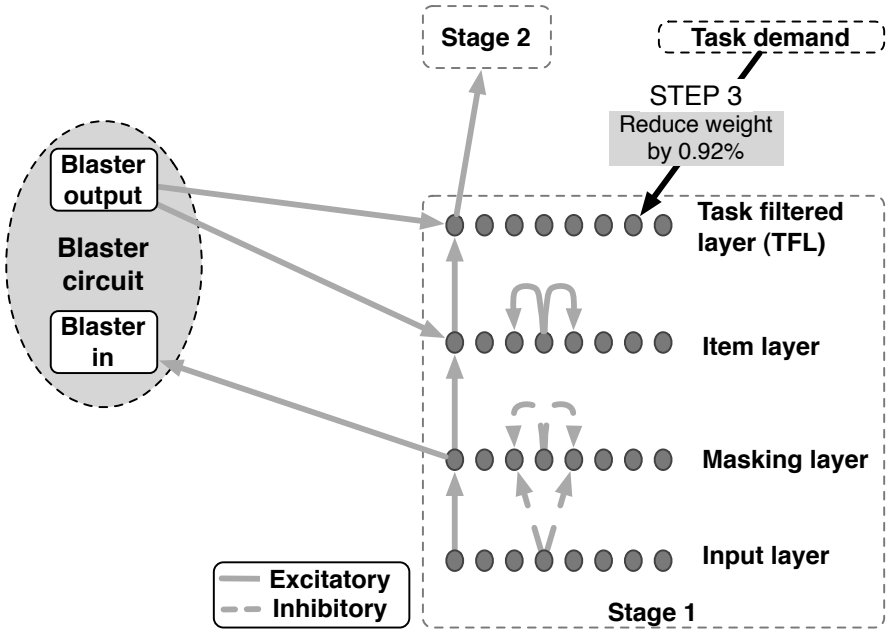


*Figure 32* Step 3 of simulating skeletal presentation. As indicated in the figure, the weight of the connection from task demand to target nodes in the TFL is reduced by 0.92%.

there is no need for the task demand mechanism to be as strong. Conceptually, in skeletal, the focus of selection moves earlier and reducing the strength of the task filter reflects this adjustment of focus. In other words, since the system can select earlier with skeletal, its later selection mechanism (at the TFL) can be more liberal. Consequently, in our second manipulation to the architecture of the $ST^2$ model, we reduce the weight from task demand to target nodes in the TFL by 0.92% of the original value (see Figure 32).

**Results** The reduction in task demand for skeletal presentation means that target nodes have less activation at the TFL. Relatively strong targets can nevertheless initiate a tokenisation process despite lower activation levels. Weak targets, however, fail to overcome the threshold for tokenisation and cannot proceed into stage two for working memory encoding.

As seen in Figure 33, weakening task demand in skeletal presentation reduces the size of the virtual P3 for skeletal targets. As the virtual P3 for skeletal targets is now considerably smaller than for RSVP targets, this qualitatively replicates the human ERP data. The skeletal virtual P3 with weaker task demand (Figure 33) also lasts slightly longer than the skeletal virtual P3 from Figure 31. This is because the reduction in task demand reduces

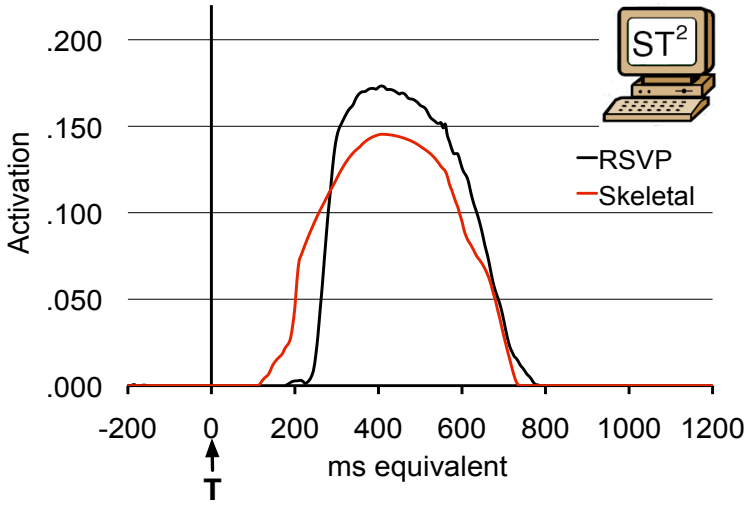


*Figure 33* After Step 3: Virtual P3 for the RSVP and skeletal condition. The RSVP vERP is baseline corrected to -200 to 0ms with respect to target onset to account for distractor related activity, which is absent in the skeletal RSVP condition. 'T' indicates the presentation of the target and ERPs time locked to presentation of the target.

the activation strength of targets in stage two. Less target activation in stage two slows down the tokenisation process, which prolongs the virtual P3 component.

After this second modification to the model's architecture, weak skeletal targets have too little activation for tokenisation and are 'missed'. The $ST^2$ model now generates a simulated accuracy of 85% for skeletal targets, which replicates the human behavioural accuracy for skeletal targets. Simulated RSVP accuracy from the model obviously remains unchanged at 77%.

### 5.3.4 Discussion

After making a few theoretically justifiable changes to the architecture of the $ST^2$ model, we are able to replicate the EEG profiles and behavioural accuracy for targets in skeletal presentation. These are compared to results from simulating target presentation in RSVP, which provides us with a potential explanation concerning the mechanisms that are causing the differences in behaviour and EEG profiles between target detection in these two paradigms.

### P1/N1 components vs. ssVEP wave

Consistent with previous findings, individually presented items in skeletal presentation produce P1/N1 early components, whereas repeatedly presented items in RSVP evoke an ssVEP wave oscillating at the frequency of stimulus presentation (see also Müller & Hillyard, 2000). We adjust the input array presented to the $ST^2$ model in that the RSVP input array contains a target embedded in a stream of distractors, whereas the skeletal input array consists of only the target and the following distractor. As a consequence, we are able replicate our human ERP results for targets in RSVP and also skeletal presentation using virtual ERPs.

In RSVP, the virtual ssVEP oscillation is caused by the stream of distractors and targets feeding into the model. Each item is presented at the input layer and propagates to the masking layer where the item experiences weak inhibition from previous items (forward masking). The item generates a short-lived peak of activation in the virtual ssVEP, before it is subject to stronger suppression from the following item in the RSVP stream (backward masking). This process repeats itself for each item in the RSVP stream and causes the oscillatory pattern that can be observed in the virtual ssVEP wave.

In skeletal presentation, the target is not forward masked, hence its activation at early layers of the $ST^2$ model is immediately larger than the activation of a target in RSVP. The following distractor inhibits the target, which causes a very short-lived dip in activation. Following this, the distractor's activation that is not suppressed through backward masking causes a large spike in the virtual ERP, which diminishes slowly according to the decay parameters of the neural network. Although visually quite different to a P1/N1 wave from the human ERP, the virtual ERP resembling early processing in the $ST^2$ model has an appropriate time course, and provides an initial qualitative fit to the human data.

### Later selection in RSVP delays working memory encoding

Our human ERP results suggest a later P3 component for targets in RSVP compared to skeletal targets. In the $ST^2$ model, tokenisation (and the virtual P3) of RSVP targets is delayed compared to skeletal targets, which is due to the earlier 'triggering' of the blaster for skeletal targets, as they can be identified as targets by visual onset. As discussed in

Section 5.3.3, the $ST^2$ model thus suggests that category distinguished RSVP tasks, where the target is embedded in a regular RSVP stream, enforce a late selection strategy. RSVP targets cannot be distinguished from distractors until they have been extensively processed, which expresses itself in delayed latency of the (virtual and human) P3 component, suggesting delayed tokenisation of RSVP targets compared to targets in skeletal presentation.

**The P2 ERP component and target selection**

Another noticeable difference in the human ERPs is the presence of a P2 wave for skeletal targets, which is absent in RSVP. Previous findings suggest that 'P2 effects occur only when the target is defined by fairly simple stimulus features, whereas P3 effects can occur for arbitrarily complex target categories' (Luck, 2005). In line with this, an analysis of previous ERP studies of the AB suggests that those studies employing colour-marked RSVP tasks, report the presence of a P2 wave in the ERP evoked by targets (Kranczioch et al., 2003; Vogel et al., 1998). Previously published ERP results from studies using RSVP tasks, where the target has to be distinguished by category, however, do not show a P2 wave (Martens, Elmallah, London, & Johnson, 2006; Martens, Munneke, et al., 2006).

If one ranks the paradigms from previously published studies by the time point after target presentation at which the task filter is engaged, this produces the following sequence. First, skeletal presentation (e.g. Ward et al., 1997), where targets appear as visual onsets and can be selected during early processing. As seen in our skeletal results from Section 5.3.2, targets in these paradigms show a distinct P2 wave. Second, colour-marked RSVP tasks (e.g. Raymond et al., 1992) in which the target's representation is fleeting due to masking from surrounding distractors, but nevertheless can be selected on the basis of visual features. The P2 for these paradigms seems slightly less distinct but still present (Kranczioch et al., 2003; Vogel et al., 1998). Third, category-distinguished RSVP tasks (e.g. Chun & Potter, 1995) where the target does not coarsely differ in terms of visual features and has to be processed for category before it can be selected. Targets in these paradigms show no P2 wave (Martens, Elmallah, et al., 2006; Martens, Munneke, et al., 2006 and our RSVP results from Section 5.3.2). Consequently, there seems to be a reciprocal relationship between the time point at which the target can be selected and size of the P2 component. In other words, the earlier the visual system is able to identify an item as

a target, the larger the P2 component of the ERP.

**Does the P2 reflect the triggering of the blaster?** The connection between the P2 and target selection is noteworthy, as - unlike its neighbouring ERP components P1, N1 and P3 - the P2 component has remained relatively unexplored. Although the ERP-AB studies mentioned earlier (Kranczioch et al., 2003; Vogel et al., 1998) found a the P2 component to be reduced in amplitude during the AB, they acknowledge that the cognitive processes underlying the P2 are relatively unknown. In fact, it is subject to debate whether the P2 reflects a perceptual or a post-perceptual process (Luck & Hillyard, 1994a; Hillyard & Münte, 1984; Kenemans, Kok, & Smulders, 1993).

We can use the $ST^2$ model to speculate that the P2 component might reflect neural activation caused by the 'triggering' of the blaster, thus indicating that the system has detected something salient. As discussed earlier, a distinct P2 component is only visible in paradigms where the blaster can be triggered from early layers of processing, i.e. skeletal presentation and (to a lesser extent) also colour-marked RSVP paradigms. According to our hypothesis, the blaster is triggered much earlier in such paradigms, namely at the masking layer. Under the assumption that the P2 component does reflect the triggering of the blaster, whereas the P3 component is the correlate of tokenisation, this suggests the following prediction:

For paradigms that allow early target selection (such as skeletal presentation or colour-marked RSVP), the P2 occurs prior to the P3 and thus the P2 is visible as a distinct component of the ERP. This is indeed what we observe in skeletal presentation and also colour-marked RSVP. In category-distinguished RSVP tasks (e.g. letters-in-digits), on the other hand, the blaster is triggered from a layer (i.e. the TFL) that is also involved in target tokenisation. Consequently, these two processes tend to coincide in time, which is why the $ST^2$ model predicts that, with late selection, the P2 and P3 components will overlay in the ERP waveform. Again, this is what we observe in the human data, where the ERP for RSVP targets contains a P3 but shows no distinct P2 component.

### Relating the P2 and N2pc components to the blaster

In the previous paragraphs, it is hypothesised that the P2 component might reflect the 'triggering' of (or input to) the blaster. The N2pc component, on the other hand, is hypothesised to reflect the firing and thus the output of the blaster (see Section 4.5.3 and also Chapter 7).

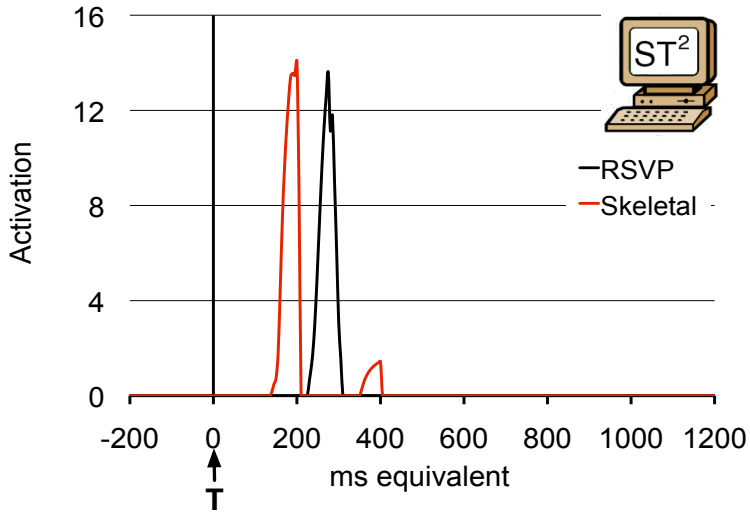


*Figure 34* Virtual N2pc component from the ST$^2$ model's blaster output nodes for RSVP and skeletal presentation. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target.

This suggests a prediction for the EEG results of an experiment investigating the N2pc component in skeletal presentation. Targets in skeletal presentation trigger the blaster at an earlier timepoint, which causes the blaster to fire earlier. Consequently, as seen in Figure 34, the virtual N2pc component occurs earlier for targets in skeletal presentation when compared to RSVP[4]. Hence, we predict an earlier N2pc component for skeletal targets when compared to targets in RSVP. To our knowledge, the N2pc component during skeletal

[4]The reader will note the small amount of activation occurring at around 400ms in the virtual ERP for skeletal presentation. This activation is generated by the distractor that is presented after the target, which also fires the blaster. As the blaster is inhibited by the tokenisation process of the target when the distractor is presented, this activation is small and stays below threshold. As discussed earlier, however, if one would want to simulate paradigms that are similar but not identical to skeletal presentation, the fact that distractors can potentially also fire the blaster might impose problems, which would have to be dealt with in future work.

presentation has not yet been investigated. Hence, this is a prediction from the $ST^2$ model that could be validated in future work.

**Behavioural prediction: No lag 1 sparing in skeletal presentation**

After modifying the $ST^2$ model, we are able simulate skeletal presentation and can qualitatively replicate the human data in our single target paradigm in terms of behavioural accuracy and virtual ERPs. Our change to the model architecture also suggests the following prediction for the behavioural results of a two target paradigm investigating the AB.

In the regular RSVP paradigm that is commonly used to investigate the AB, the targets are embedded in a continuous stream of distractors. If one presents two targets (T1 & T2) in immediate succession and they are followed by at least one distractor to ensure backward masking, observers are likely to report T2 correctly. In fact, T2 accuracy is often higher than accuracy at detecting a single target. This effect is called lag 1 sparing (see Section 2.3.2). In an AB experiment using the skeletal paradigm, however, we would expect there to be no such second target advantage.

In a regular RSVP paradigm - according to the $ST^2$ model - T1 triggers the blaster when it reaches the TFL. However, there is some temporal delay between the blaster being triggered and the timepoint of its full effect on the item layer and the TFL. In regular RSVP, this means that a T2 appearing at the TFL shortly after T1 will get much of the benefit of T1's blaster response. This results in the increased accuracy at detecting T2, as exemplified by lag 1 sparing.

In skeletal presentation, however, T1 triggers the blaster at an earlier timepoint, i.e. as soon as T1 reaches the masking layer. Despite there being some temporal delay until the blaster becomes fully active, the blaster will have its major effect by the time T1's activation has reached the item layer and the TFL. As skeletal presentation causes the whole activation profile of the blaster to be shifted back in time, the blaster is no longer active when T2 arrives, as it is already being suppressed by T1's tokenisation. Hence, the $ST^2$ model's prediction for skeletal presentation is that T2 detection accuracy at lag 1 should be low, i.e. T2 should be 'blinked' instead of there being lag 1 sparing.

To reiterate, the $ST^2$ model predicts that the time point at which T1 fires the blaster

(earlier for skeletal than RSVP targets) has an effect on T2's accuracy, i.e. lag 1 sparing only obtains for RSVP targets but not in skeletal presentation. If one analyses the behavioural data from AB studies that employed the skeletal paradigm (as depicted in Figure 27), we see that exactly this is the case[5]. In regular RSVP, if T2 is presented immediately following T1, its accuracy will be excellent, i.e. we observe lag 1 sparing (Figure 7). In skeletal presentation, however, there is no lag 1 sparing. In fact, T2 accuracy at lag 1 is the lowest point of the AB (see Figure 27).

**Is skeletal presentation an equal substitute for RSVP?**

Aside from theoretical considerations, the regular RSVP paradigm has a number of practical disadvantages. Due to the fast presentation rate, regular RSVP streams contain a relatively large number of distractors, so the typical total length of an RSVP stream is around 2-3 seconds. Furthermore, the rapid presentation of items is often taxing for participants, especially when they are participating in a longer experiment. The 'length of the stream' issue is particularly important when conducting EEG or magnetoencephalography (MEG) experiments. In order to increase the signal-to-noise ratio through averaging, each condition is presented several times to the participant. The relatively long duration of an RSVP stream compared to the presentation of a single target is troublesome, as it inflates the duration of the experiment. As experimental time in an EEG/MEG laboratory is costly, there is a major incentive to keep the duration of an experiment as short as possible.

The 'skeletal RSVP task minimises demands both on selective attentional processing and on location switching mechanisms' (Ward et al., 1997), while nevertheless seeming to reveal the attentional limitations underlying the AB. Thus, due to less complexity and reduced duration of experiments through shorter streams, skeletal presentation seems ideal for studies employing MEG or EEG to study the AB. As a recent study investigating the AB by means of MEG and the skeletal paradigm states: 'an AB effect is observed whether targets are embedded in a 20-item RSVP stream or just presented on their own followed by masks (Duncan et al., 1994; Ward et al., 1997). In order to save measurement time, we

[5]Note that - unlike the Duncan et al. (1994) study described in earlier sections of this chapter - the targets in these studies were not spatially offset. Hence, the lack of lag 1 sparing is not due to a move in spatial location (see Visser et al., 1999) but seems to be solely due to skeletal presentation.

decided to employ this abbreviated version for our study' (Kessler et al., 2005b).

From the results presented in this chapter, however, we can argue that there are considerable differences in how targets are processed depending on whether the paradigm employs skeletal presentation or RSVP. Consequently, EEG/MEG data collected from an AB experiment employing the skeletal paradigm may not be directly comparable to EEG/MEG data from experiments using regular RSVP and should be interpreted with caution.

## 5.4 Removing the T+1 distractor

In line with investigating the role of distractors for target processing in RSVP, we also explore the effect of removing the distractor following the target. After removing forward masking by means of the skeletal paradigm in the previous section, the 'T+1 blank' paradigm allows us to test the influence of reduced backward masking on behavioural and electrophysiological correlates of target detection. In order to simulate T+1 blank streams in the $ST^2$ model, the input value of the distractor following the target is set to zero. Compared to regular RSVP, the architecture of the $ST^2$ model is not modified.

### 5.4.1 Results

#### Behaviour

The removal of the distractor following the target has a positive effect on target accuracy. Participants do better at detecting targets in the RSVP T+1 blank condition (85%, SEM 4) as compared to detecting targets that are surrounded by distractors in a regular RSVP stream (72%, SEM 4). The difference is significant, $F(1,19) = 111.4$, $MSE < 0.01$, $p < 0.001$. The $ST^2$ model produces a qualitative fit of human behavioural accuracy for single targets in the T+1 blank ($ST^2$ accuracy: 100%) and RSVP ($ST^2$ accuracy: 77%) condition.

#### Electrophysiology

**Human ERP** Figure 35A illustrates how the P3 component occurs later for targets in RSVP compared to T+1 blank targets (300-1050ms, RSVP: mean 500ms (SEM 19) vs. T+1 blank: mean 472ms (SEM 16); $F(1,19) = 4.5$, $MSE = 1647$, $p = 0.048$). Furthermore,

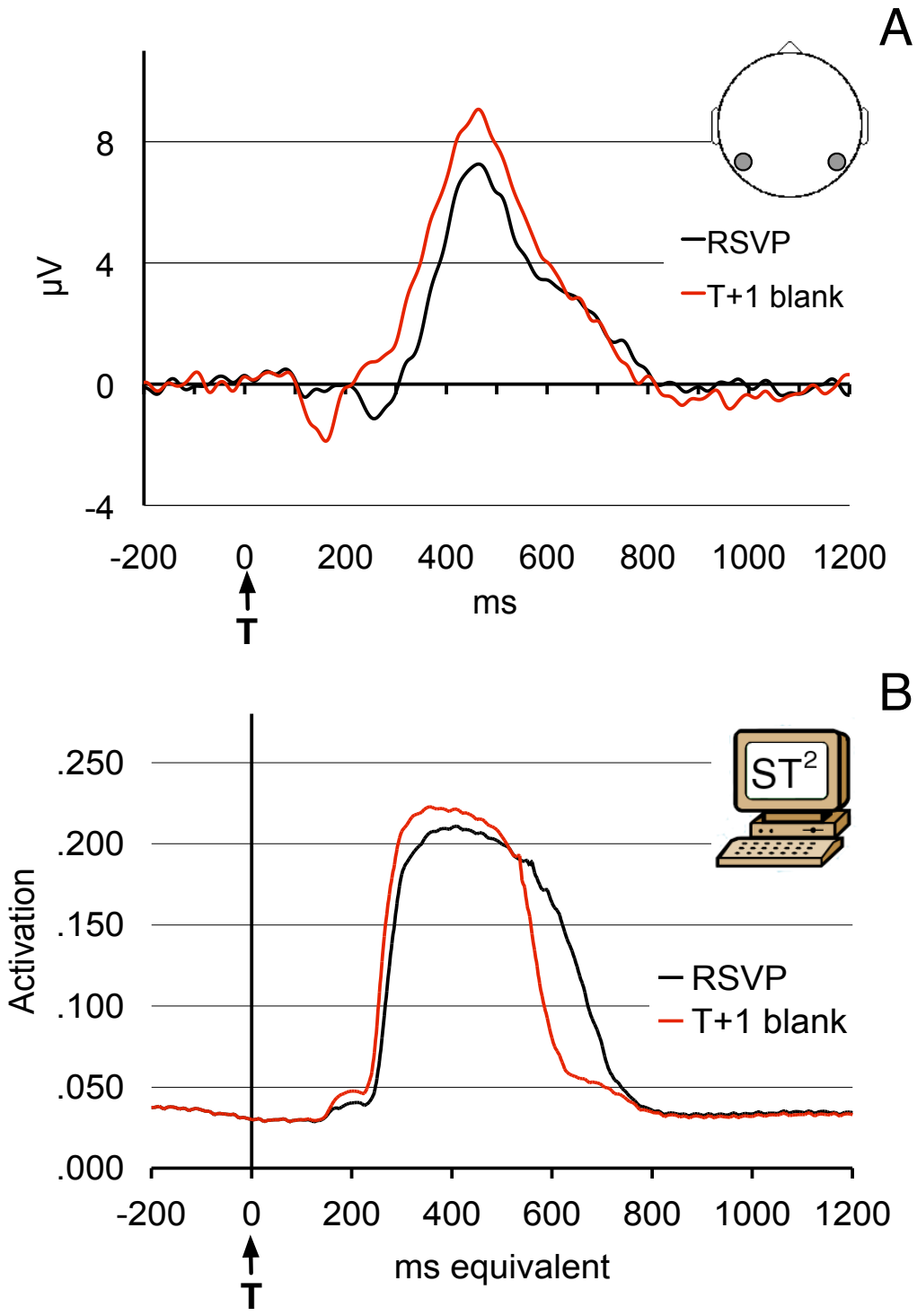


*Figure 35* Panel A: Human ERP for targets in the RSVP and T+1 blank conditions averaged across P7 and P8 electrode locations. Positive is plotted upwards. Panel B: Virtual P3 from the $ST^2$ model for the RSVP and T+1 blank conditions. 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target.

the P3 is smaller for standard RSVP targets (mean value 2.3$\mu$V, SEM 0.4) than for targets followed by a blank (mean value 2.8$\mu$V, SEM 0.4); F(1,19) = 4.7, MSE = 0.5, p = 0.043.

**Virtual ERP** In the ST$^2$ model, backward masking is simulated by means of inhibition at the masking layer. If multiple items are active together at the masking layer, as is the case with targets and distractors in RSVP, they will inhibit each other. If a target is following by a blank interval, the target is inhibited less strongly than a target that is embedded in a continuous stream of distractors. A less strongly masked target consequently has more activation when it reaches the TFL at the end of stage one. Swifter tokenisation with more activation is reflected in a larger and earlier virtual P3 component for T+1 blank targets compared to targets in RSVP, as seen in Figure 35B.

### 5.4.2 Discussion

#### Bottom-up strength: T+1 blank & target difficulty

The ST$^2$ model replicates the human EEG data because T+1 blank targets are able to initiate a stronger tokenisation than targets in standard RSVP. This causes a larger virtual P3 component, which replicates the effect observed in the human P3 component. It also means that T+1 blank targets are more likely to be 'seen' by the model, which provides a qualitative match of the human accuracy. The ST$^2$ model thus proposes that reduced backward masking and making targets easier to detect due to intrinsic stimulus characteristics (as seen in the easy/hard analysis of Chapter 6) have a similar effect, namely an increase in bottom-up trace strength.

#### Reduced masking leads to earlier tokenisation

More activation in terms of bottom-up strength of T+1 blank targets at the TFL leads to an earlier blaster response. In consequence, these targets are able to initiate tokenisation at an earlier time point. Hence, the virtual P3 replicates an earlier human P3 latency for T+1 blank targets. The ST$^2$ model thus proposes that the differences in behavioural accuracy scores and EEG signatures of targets presented in a T+1 blank stream compared to regular RSVP occur due to differences in bottom-up trace strength.

In contrast, a recent MEG study investigating the effect of masking on targets during the AB (Kessler et al., 2005b), reported an earlier M300 (MEG P3 equivalent) latency in the prefrontal MEG source cluster for targets followed by a mask compared to targets followed by a blank. The authors conclude that the mask *accelerates* target processing, which is the opposite of the ERP effect reported in this chapter and in contrast with the theory underlying the $ST^2$ model. However, in Kessler et al. (2005b), the MEG M300 component was extracted from a prefrontal brain area, whereas the EEG P3 component is commonly found at parietal sites. Hence, Kessler et al. (2005b)'s MEG results cannot be directly compared to the P3 component analysis from the EEG results presented here.

## 5.5 Conclusion

This chapter contributes to the theme of this thesis in terms of informing the theoretical discussion of target processing during high temporal demands. We modify the $ST^2$ model to simulate skeletal presentation and demonstrate how the virtual ERP technique can be used to explore the ERP effects observed when investigating the role of distractors for target processing in RSVP.

#### Bottom-up strength vs. effort influencing the P3 component

Our results suggest that in each of the paradigms presented in this chapter, the P3 component is affected in a different way. In skeletal presentation, the lack of a continuous stream of distractors causes the human P3 to be smaller compared to targets in regular RSVP. If the distractor following the target in RSVP is omitted, this increases the size of the P3 compared to targets in regular RSVP.

The ERP results reported in this chapter contribute to the discussion of the meaning of P3 amplitude for targets in RSVP (see Chapter 6 for a further investigation of this issue). In skeletal presentation, the difference in P3 amplitude might be due to the amount of effort subjects allocate towards the processing of the target. Due to blocked design, participants knew before the stream started when to expect a target in skeletal or in RSVP, respectively. It is harder to distinguish targets from surrounding distractors in RSVP than detecting targets in skeletal presentation, which is reflected by the difference in behavioural accuracy

scores between the two conditions. If subjects have a previous indication of how difficult the following task is going to be, the amount of preallocated effort (and not bottom-up strength) becomes the main modulator of P3 amplitude (Sirevaag, Kramer, Coles, & Donchin, 1989; Wickens, Kramer, Vanasse, & Donchin, 1983; Kramer & Hahn, 1995). Consequently, RSVP targets have larger P3 components than targets in skeletal, because - on average - subjects decide to allocate more resource to the processing of a target in RSVP compared to a target in skeletal presentation.

RSVP and T+1 blank trials, however, were presented in an intermixed design and thus participants could not predict the occurrence of each condition. Hence, P3 amplitude cannot be a correlate of preallocated effort. We argue that, instead, the reduction in backward masking modulates intrinsic target strength, similar to the easy/hard manipulation that is discussed in Chapter 6. T+1 blank targets are masked less strongly, which means they have more bottom-up strength leading to a stronger tokenisation process. As indicated by the virtual ERP from the $ST^2$ model, there is a positive correlation between increased bottom-up target strength and larger P3 amplitude, which is in line with intrinsically easier targets having a larger P3 (Johnson, 1986; Kok, 2001).

**Single target findings benefit AB studies**

The insights gained from our EEG results are important for interpreting the data from studies investigating the AB. As our paradigm employed just a single target, we can assume that the observed ERP effects are solely due to the presence/absence of distractors in each of the paradigms. Because of the close temporal proximity between targets during the AB, such a prerequisite cannot be assumed when employing two-target paradigms. This can be a problem when investigating the meaning of P3 amplitude for targets in RSVP (see Chapter 6). Our single target study, however, suggests how strategic resource allocation and bottom-up strength have opposite effects on P3 amplitude and indicates the circumstances under which each of these factors become the dominating factors influencing the P3 component. As we will further discuss in Chapter 6, such knowledge is critical for formulating theories of the AB.

# Chapter 6

# The attentional blink reveals serial working memory encoding

In this chapter, we evaluate the resource sharing hypothesis and the $ST^2$ model as two competing theories of the AB. The resource sharing hypothesis proposes a dynamic distribution of resources over a time span of up to 600ms during the AB. The $ST^2$ model, on the other hand, argues that, due to serial working memory encoding, targets are encoded in separate episodes during the AB and that, due to joint consolidation, lag 1 is the only case were resources are shared. We use the $ST^2$ model to generate predictions by means of virtual ERPs for each of the conditions of interest. In a first analysis, we investigate the meaning of P3 component amplitude evoked by targets in RSVP. The results suggest that, at least in this context, P3 amplitude is an indication of bottom-up strength, rather than a measure of cognitive resource allocation. Second, our results suggests that T1 consolidation is not affected by the presentation of T2 during the AB. However, if targets are presented in immediate succession (lag 1 sparing), they are jointly encoded into working memory. The EEG results are in line with the virtual ERP predictions and thus support the theory underlying the $ST^2$ model.

## 6.1 Introduction

In daily life, humans have to cope with an environment consisting of simultaneously occurring events and concurrent sensory input. In order to survive in this parallel world, attention allows us to filter out irrelevant information. On the one hand, attention lets us focus on one task at a time, while on the other hand, we are also often able to perform multiple tasks simultaneously. Thus it seems that cognitive resources can be shared between tasks, suggesting a notion of divided attention. This distribution of attention, however, seems to come with concomitant costs and limitations both in terms of performance accuracy and reaction times. In this chapter, we investigate the extent to which attentional resources can be shared over time and the cost associated with it.

In spatial attention, the visual system was long assumed to operate in a serial manner, in that it was restricted to selecting information from only one location at a time. Attention was considered to move through the visual field in the form of a single spotlight (von Helmholtz, 1867; Broadbent, 1958; Posner, 1980; Eriksen & Yeh, 1985). However, Pylyshyn and Storm (1988), amongst others, disproved these classical theories by showing that humans are capable of simultaneously tracking multiple objects in space. Some of the new theories preserve the idea of a single focus of attention, which sequentially switches between targets (Pylyshyn & Storm, 1988; Oksama & Hyöna, 2004); others propose a notion of concurrent multifocal attention, which can be focused on more than one location at a time (Castiello & Umilta, 1992; Awh & Pashler, 2000; McMains & Somers, 2004).

Whether attention is a single spotlight, switching rapidly between locations, or whether attentional resources are distributed across multiple locations, simultaneous perception of multiple objects in space requires some notion of resource sharing. In line with this argument, Cavanagh and Alvarez (2005) conclude that the 'trade-off between capacity and feature encoding (Oksama & Hyöna, 2004; Bahrami, 2003; Saiki, 2003) suggests that attention has a fixed total bandwidth for selection and the bandwidth can be shared across several input channels or targets'. Hence, although the system is capable of tracking multiple objects at a time, there is a fixed amount of attentional resource. As this resource is shared across increasing numbers of targets, overall performance at the task decreases.

Recently, it has been proposed that the notion of a shared attentional resource with fixed

capacity could be extended to the temporal domain (Shapiro et al., 2006). Accordingly, if multiple target items are presented at the same spatial location within a very short period of time, the system allocates a certain amount of the resource to each of the targets and they are, at least to some extent, processed in a concurrent manner. Hence, if one of the targets is processed more extensively, less resource is available for other targets, which has a detrimental effect on target detection accuracy, thus providing an explanation for the AB.

### 6.1.1 Resource sharing vs. two-stage theories

As discussed in Section 3.1.3, the resource sharing hypothesis suggests that the AB is an artifact of compromised allocation of attention (Shapiro et al., 2006). If the system allocates less resource to T1, more attention is available for T2 and T2 is more likely to be detected. If, however, 'too much' resource is allocated to T1, T2 is more likely to be missed, which results in an AB (Kranczioch et al., 2007).

In contrast and as discussed in Section 3.1.2, two-stage theories (Chun & Potter, 1995) propose that the AB reveals a cognitive mechanism, which ensures serial working memory encoding to protect the integrity of an attentional episode (Wyble et al., 2009). If T2 is presented during the AB window, its working memory consolidation is delayed until T1 has been successfully encoded. At lag 1, however, this 'protection mechanism' breaks down and T1 and T2 are encoded into a single attentional episode. Joint encoding increases T2 accuracy at lag 1, but comes at the cost of increased swaps (i.e. T1 and T2 are identified correctly but reported in the wrong order) and reduced T1 accuracy (see Section 2.3.2).

### 6.1.2 The P3 component as a measure of resource allocation?

The resource sharing hypothesis was formulated in response to a number of findings derived from EEG (Martens, Elmallah, et al., 2006; Kranczioch et al., 2007) and MEG (Shapiro et al., 2006) experiments investigating the AB. These authors base their argument on the assumption that the size of the P3 component evoked by a target in RSVP reflects the amount of resources invested into processing that target.

However, in an extensive review of the P3 component, Kok (2001) comes to the conclusion that 'the sensitivity of P3 amplitude as a measure of processing capacity has only been

convincingly demonstrated in a restricted number of studies in which capacity allocation was under voluntary control, and the structural characteristics of the task (e.g. task complexity, perceptual quality of the stimuli) did not change'. Accordingly, P3 size increases if observers know *beforehand* that the task is going to be harder, and allocate more cognitive resource to it (Sirevaag et al., 1989; Wickens et al., 1983; Kramer & Hahn, 1995). When task difficulty is determined only by intrinsic stimulus properties, however, there is a reciprocal relationship between increasing task difficulty and P3 amplitude (Johnson, 1986).

This distinction is critical when using the P3 component to evaluate theories of the AB. In category-distinguished AB tasks, target items are often letters presented in a stream of digit distractors (see Section 2.3.1 and also Figure 6A for a depiction of this AB task). Due to their shape, some target letters are masked more strongly by the distractors than others, thus target letters can be categorised by their individual accuracy scores, yielding a measure of task difficulty according to intrinsic stimulus properties. We will use the terms 'easy' and 'hard' to categorise letters according to their individual accuracy scores. In RSVP, target letters commonly appear in random order. As observers cannot predict whether an upcoming target in RSVP will be easy or hard, they do not know beforehand how much resource to allocate to the target. Hence, in Kok (2001)'s terms, resource allocation is not 'under voluntary control' whereas the 'structural characteristics' of the stimuli do change and, thus, the P3 should not serve as a measure of resource allocation.

Recent articles arguing in favour of resource sharing have proposed that the allocation of resource to targets in RSVP might be random, varying from trial to trial (Shapiro et al., 2006; Kranczioch et al., 2007). If by chance more resource is allocated to T1, less attention is available for T2, thus suggesting a trade-off in accuracy and P3 sizes. Depending on how one interprets this argument of random allocation of resources, we can make two predictions about the resulting nature of P3 for easy and hard targets: (a) If resource allocation is truly random, it should produce no difference in the average P3 amplitude between easy and hard targets. (b) Alternatively, if hard targets are somehow able to instantaneously attract more resources, we should expect to observe a larger P3 for intrinsically hard targets, when compared to easy ones.

The $ST^2$ model, in contrast, makes a different prediction regarding the effects of target

difficulty, in that the amplitude of the P3 for targets in RSVP should be mainly modulated by bottom-up strength. If a target is easier to perceive due to its intrinsic stimulus characteristics, for instance if it is less strongly masked, the target has more bottom-up target strength, which leads to a larger P3 (see Section 5.4). Vice versa, a target that is intrinsically harder to detect will have less bottom-up strength, thus evoking a smaller P3 component.

### 6.1.3 Overview

In this chapter, we evaluate the resource sharing theory and the $ST^2$ model as two competing explanations of the AB. We use the $ST^2$ model's neural network implementation to generate virtual ERPs (vERPs) and compare these to human ERP data (hERPs).

To this end, we first address the question of understanding P3 amplitude differences for RSVP targets, which is critical for interpreting EEG/MEG results. Does a large P3 indicate that more effort was dedicated to the task because it was harder? Or is P3 size mainly modulated by intrinsic stimulus characteristics, in which case a larger P3 indicates that the target was particularly strong and hence easy to perceive? This question is addressed by analysing the EEG response of a single target in RSVP (from Experiment 1, see the methods section in the appendix for further details). Whether a target letter is easy (or hard) depends solely on intrinsic stimulus characteristics and, thus, the hERP data (and corresponding vERPs from the $ST^2$ model) can be used to evaluate the competing hypotheses of P3 amplitude described in the previous section.

We then analyse target-related EEG activity in a two-target AB paradigm (from Experiment 2, see the methods section in the appendix for further details) and use the $ST^2$ model to generate corresponding vERPs. Although the resource sharing theory lacks a clear formal description, it does make a key prediction for EEG/MEG data. The resource sharing theory suggests that targets indirectly compete during the AB through the amount of resources allocated to each of them. Hence, it predicts that the T1 P3 component should be larger for trials in which T2 is missed during the AB, as too much resource was invested into the processing of T1. On the other hand, if T2 is seen during the AB, the T1 P3 is likely to be smaller as subjects were able to allocate resource more evenly between targets. In contrast, the $ST^2$ model proposes that targets are encoded one at a time, thus

emphasising the serial nature of working memory encoding during the AB. This suggests the following prediction for the EEG/MEG correlates of target encoding during the AB. T1 consolidation (as exemplified by T1's P3 component) should influence T2 processing in both behavioural and electrophysiological terms, since T2s have to 'wait' until T1 has been consolidated. The reverse, however, is not the case, i.e. T1's P3 should be unaffected, regardless of whether T2 is seen or missed and thus the influence between T1 and T2 is unidirectional. Only if the targets appear in immediate succession, as is the case at lag 1, can there be mutual interference.

## 6.2 Methods

This chapter is based on behavioural and EEG data from Experiment 1 and 2. Please refer to the appendix for a detailed overview of the methods employed in these experiments. Details specific to the analyses presented in this chapter are described in the following sections.

### 6.2.1 Experiment 1

#### EEG analysis

ERPs were time locked to the presentation of the target and extracted from -200 to 1200ms with respect to target presentation. After artifact rejection, ERPs in each of the conditions ('easy' and 'hard') contained 912 and 662 epochs, respectively. Activity from the Pz (midline parietal) electrode was used to analyse the P3 component. Since only seen targets evoke a P3, while missed targets do not (e.g. Kranczioch et al., 2003), ERPs were generated only from trials in which the target was correctly identified.

### 6.2.2 Experiment 2

#### EEG analysis

ERPs were time locked to T1 and extracted from -200 to 1800ms with respect to T1 onset. After artifact rejection, ERPs for each of the conditions contained the following number of trials: Lag 3 noAB - 863 epochs; Lag 3 AB - 702 epochs; Lag 8 - 1201 epochs; Lag 1 -

946 epochs. In the following sections, 'Lag 3 noAB' refers to the conditions when T2 was presented at lag 3 and both targets were correctly identified so that an attentional blink did not occur. 'Lag 3 AB' is the condition when T1 was accurately reported but T2 could not be correctly identified and hence the observer experienced an attentional blink on that particular trial. The 'Lag 8' and 'Lag 1' conditions describe scenarios in which T2 was presented at the given lag (with respect to T1) and both targets were correctly reported. Experiment 2 contained a bilateral RSVP paradigm as we also investigate the lateralised N2pc component during the AB (see Chapter 7). Target presentation to the left and right of fixation was equally probable, randomised and the P3 was recorded from the midline Pz electrode. Hence, bilateral presentation was irrelevant for the purpose of the analyses in this chapter.

## 6.3 Results: Experiment 1

### 6.3.1 Behaviour

We determine the accuracy score for each target letter by using the behavioural results for T1 accuracy per letter from a previously published study (Bowman & Wyble, 2007), which employed a similar RSVP paradigm [1]. Accordingly, all targets are classified as belonging either to the 'easy' or the 'hard' group of target letters. By dividing targets a priori (with respect to the experiment reported here), we counter arguments that our subdivision into easy and hard reflects random variation in attentional state (i.e. alertness) of subjects, rather than fluctuations in intrinsic stimulus strength. The fact that it is the same letters that are easy (respectively hard) in the Bowman and Wyble (2007) experiment and the experiment reported here is strong evidence that variation in intrinsic stimulus characteristics underlies this subdivision.

The behavioural results from Experiment 1 show that the 'hard' target letters (E, C, B, P, F, J and R) have an average accuracy of 62% (SEM 4), whereas the 'easy' targets (T, K, U, V, L, D and G) have an average accuracy of 82% (SEM 4). The difference in

[1] The 54ms SOA experiment from Bowman and Wyble (2007) also used a presentation rate of approx. 20 items per second and the resulting T1 accuracy (averaged across conditions where T2 is presented at lag 12/648ms, lag 14/756ms and lag 16/864ms) is comparable to the accuracy of detecting single targets in Experiment 1 (72% vs. 77%).

accuracy scores between the easy and the hard target group is highly significant ($F(1,19) = 94.1$, $MSE < 0.01$, $p < 0.001$).

In the $ST^2$ model, a target is classified as hard if its strength value is less than or equal to the value of distractors (strength values 0.442 to 0.526). Target values above those of distractors contribute to the easy condition (strength values 0.540 to 0.610). The $ST^2$ model provides a qualitative fit of the behavioural accuracy scores for the hard ($ST^2$ accuracy: 57%) and easy ($ST^2$ accuracy: 100%) conditions.

### 6.3.2 Human ERP

As seen in Figure 36A, the P3 for easy targets has a significantly larger amplitude than the P3 for hard targets ($F(1,19) = 5.3$, $MSE = 4.3$, $p = 0.033$). The mean amplitude in the 300-600ms post-target area is 8.3$\mu$V (SEM 1.1) for easy targets and 6.8$\mu$V (SEM 0.9) for hard targets. Although the P3 for hard targets starts slightly later than the P3 for easy targets, it also returns back to baseline more rapidly and thus the small difference in 50% area latency analysis (Luck & Hillyard, 1990) is non-significant (easy targets: mean 447ms (SEM 12) vs. hard targets: mean 455ms (SEM 10); $F(1,19) = 0.8$, $MSE = 763$, $p = 0.371$).

### 6.3.3 Virtual ERP

In the $ST^2$ model, easy targets have higher input strength and thus generate more activation than hard targets. Figure 36B illustrates how the vP3 is larger in amplitude for easy compared to hard targets (mean vP3 amplitude: Easy 0.203 vs. Hard 0.189). Once target activation reaches later parts of stage one, easy targets trigger an earlier blaster response, which causes these items to be encoded into working memory more rapidly. The result is a slightly earlier vP3 component for easy (vP3 50% area latency: 455ms equivalent) compared to hard (vP3 50% area latency: 460ms equivalent) targets, as seen in Figure 36B.

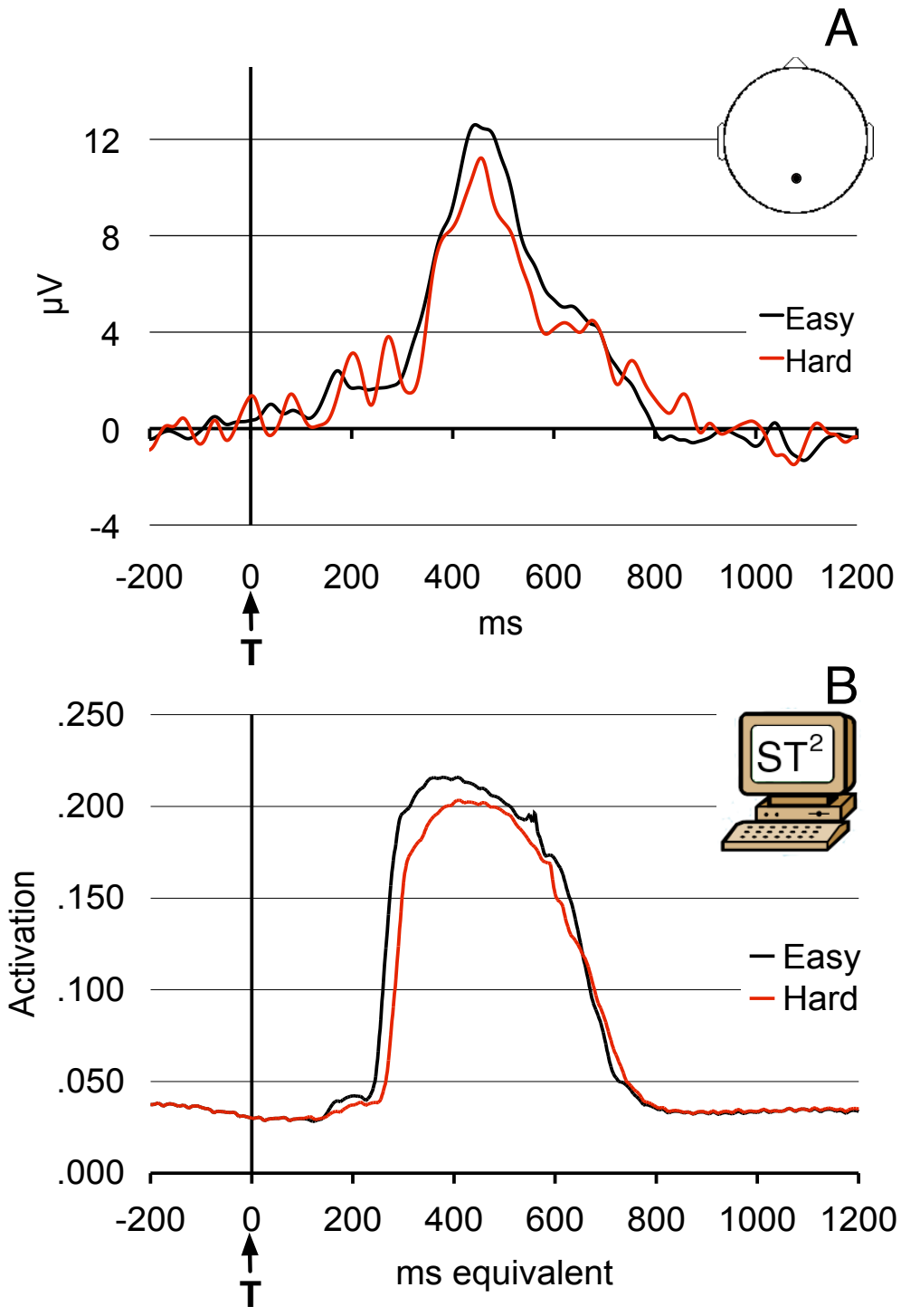


*Figure 36* Panel A: hERP P3 component from Pz for the easy and hard condition. Positive is plotted upwards. Panel B: $ST^2$'s vERP containing the virtual P3 component for the easy and hard condition. For both panels, 'T' indicates the presentation of the target and ERPs are time locked to presentation of the target.

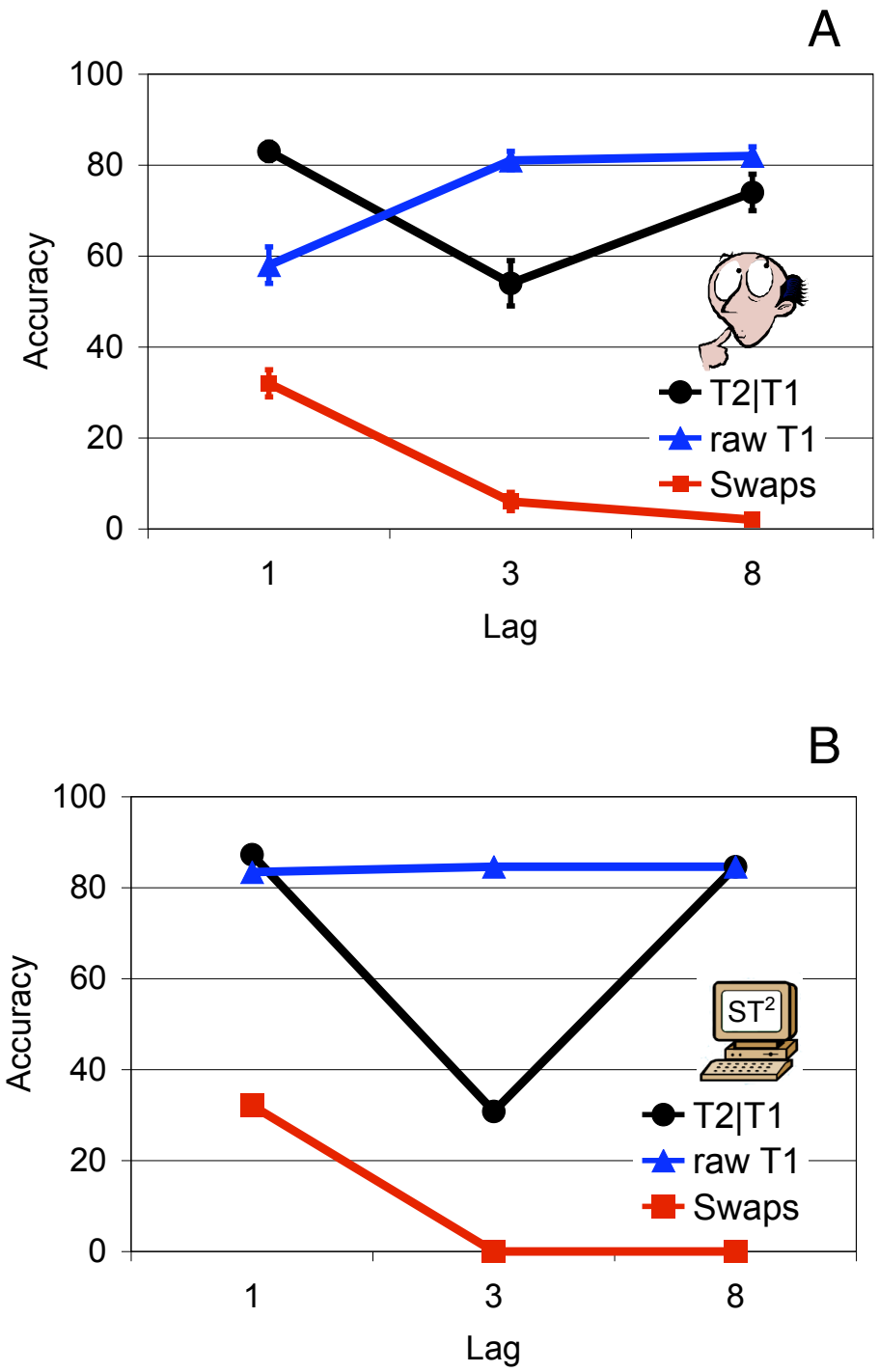


*Figure 37* Panel A: Human behavioural accuracy data for lag 1, lag 3 and lag 8. Panel B: Simulated behavioural accuracy of the $ST^2$ model for lag 1, lag 3 and lag 8. Circles indicate T2 accuracy conditional on correct T1 report, triangles represent raw T1 accuracy and squares indicate swaps, i.e. the condition when T1 and T2 were correctly identified but reported in the wrong order. In panel A, error bars represent standard error of the mean.

## 6.4 Results: Experiment 2

### 6.4.1 Behaviour

#### Attentional blink

As shown in Figure 37A, human accuracy at identifying T2 (conditional on correct report of T1) shows a significant effect of lag ($F(2,17) = 15.58$, MSE = 0.03, GG-$\varepsilon$ = .74, $p < 0.001$). Pairwise comparisons emphasise the presence of an AB. T2 accuracy is significantly lower at lag 3 compared to lag 8 ($F(1,17) = 11.66$, MSE = .03, $p = .003$) and lag 1 ($F(1,17) = 60.88$, MSE = 0.01, $p < 0.001$). If T2 is presented in immediate succession to T1 (lag 1), T2 accuracy is significantly higher than T2 accuracy at lag 8 ($F(1,17) = 5.41$, MSE = 0.01, $p = 0.033$). As seen in Figure 37B, the $ST^2$ model replicates a U-shaped AB curve. T2 accuracy (conditional on correct report of T1) is reduced at lag 3 compared to lag 8 and lag 1. Furthermore, T2 accuracy at lag 1 is slightly higher than at lag 8.

When comparing the simulated accuracy of the $ST^2$ model to the behavioural data from Experiment 2, it should be noted that the model was originally configured to replicate the AB curve published in Chun and Potter (1995). Subsequent studies (including the behavioural data from Experiment 2 as reported here) mostly reported higher lag 3 accuracy and thus a less drastic AB effect. To comply with the philosophy of changing as few parameters as possible compared to the $ST^2$ model published in Bowman and Wyble (2007), we sacrifice a perfect quantitative fit of the data from Experiment 2 and, instead, emphasise the qualitative replication of an AB effect.

#### Reduced T1 accuracy at lag 1

As depicted in Figure 37, observers are significantly worse at reporting T1 if T2 is presented at lag 1 compared to when T2 is presented at lag 3 ($F(1,17) = 49.68$, MSE = 0.01, $p < 0.001$) or lag 8 ($F(1,17) = 61.21$, MSE = 0.01, $p < 0.001$). Although this is admittedly a weak effect, the $ST^2$ model qualitatively replicates a reduction in T1 accuracy at lag 1 (Simulated T1 accuracy: lag 1 - 83%, lag 3 - 85%, lag 8 - 85%).

#### No effect on T1 accuracy when T2 is at lag 3 or 8

We observe no significant difference in T1 accuracy between T2 presented at lag 3 or lag 8 ($F(1,17) = 0.44$, MSE $< 0.01$, $p = 0.515$; see Figure 37). Furthermore, there is no difference in T1 accuracy whether an AB occurs or not (T1 accuracy conditional on seen T2 at lag 3: 79%, SEM 4; T1 accuracy conditional on missed T2 at lag 3: 78%, SEM 3; $F(1,17) = 0.03$, MSE = 0.02, $p = 0.862$). The $ST^2$ model replicates these effects, since simulated T1 accuracy is at baseline irrespective of whether T2 is presented at lag 3 or lag 8.

#### Increased number of swaps at lag 1

Figure 37 shows that, at lag 1, we observe a high percentage of swaps, but swaps are negligible at lags 3 and 8. The difference in swaps between lag 1 and lag 3 ($F(1,17) = 58.67$, MSE = 0.01, $p < 0.001$) and also lag 1 compared to lag 8 ($F(1,17) = 133.31$, MSE = 0.01, $p < 0.001$) is highly significant. The $ST^2$ model replicates this effect, and produces a high proportion of swaps if T2 is presented at lag 1 but produces no order inversions at lags 3 and 8.

### 6.4.2 Human ERPs

Our results suggest no significant difference in mean amplitude of T1's P3 (300-600ms) with respect to T2 presentation (Figure 38A). First, there is no significant difference in T1 P3 amplitude whether an AB occurs or not (Lag 3 AB: 6.5$\mu$V (SEM 0.6) vs. Lag 3 noAB: 7.3$\mu$V (SEM 0.6); $F(1,17) = 1.91$, MSE = 2.7, $p = 0.185$). Second, there is no significant difference in T1 P3 amplitude whether T2 is presented at lag 3 or lag 8 (Lag 3 noAB: 7.3$\mu$V (SEM 0.6) vs. Lag 8: 7.0$\mu$V (SEM 0.6); $F(1,17) = 0.32$, MSE = 2.0, $p = 0.576$).

As suggested by Figure 38A, T1 P3 50% area latency (calculated for the 300-600ms window) seems to be independent of T2 presentation. First, there is no significant difference in T1 P3 latency whether an AB occurs or not (Lag 3 AB: 453ms (SEM 5) vs. Lag 3 noAB: 452ms (SEM 5); $F(1,17) = 0.02$, MSE = 241.8, $p = 0.883$). Second, whether T2 is presented at lag 3 or lag 8 has no significant effect on T1 P3 latency (Lag 3 noAB: 452ms (SEM 5) vs. Lag 8: 454ms (SEM 3); $F(1,17) = 0.18$, MSE = 191.9, $p = 0.670$).

We replicate the finding that T2 evokes a P3 component in those trials in which an AB

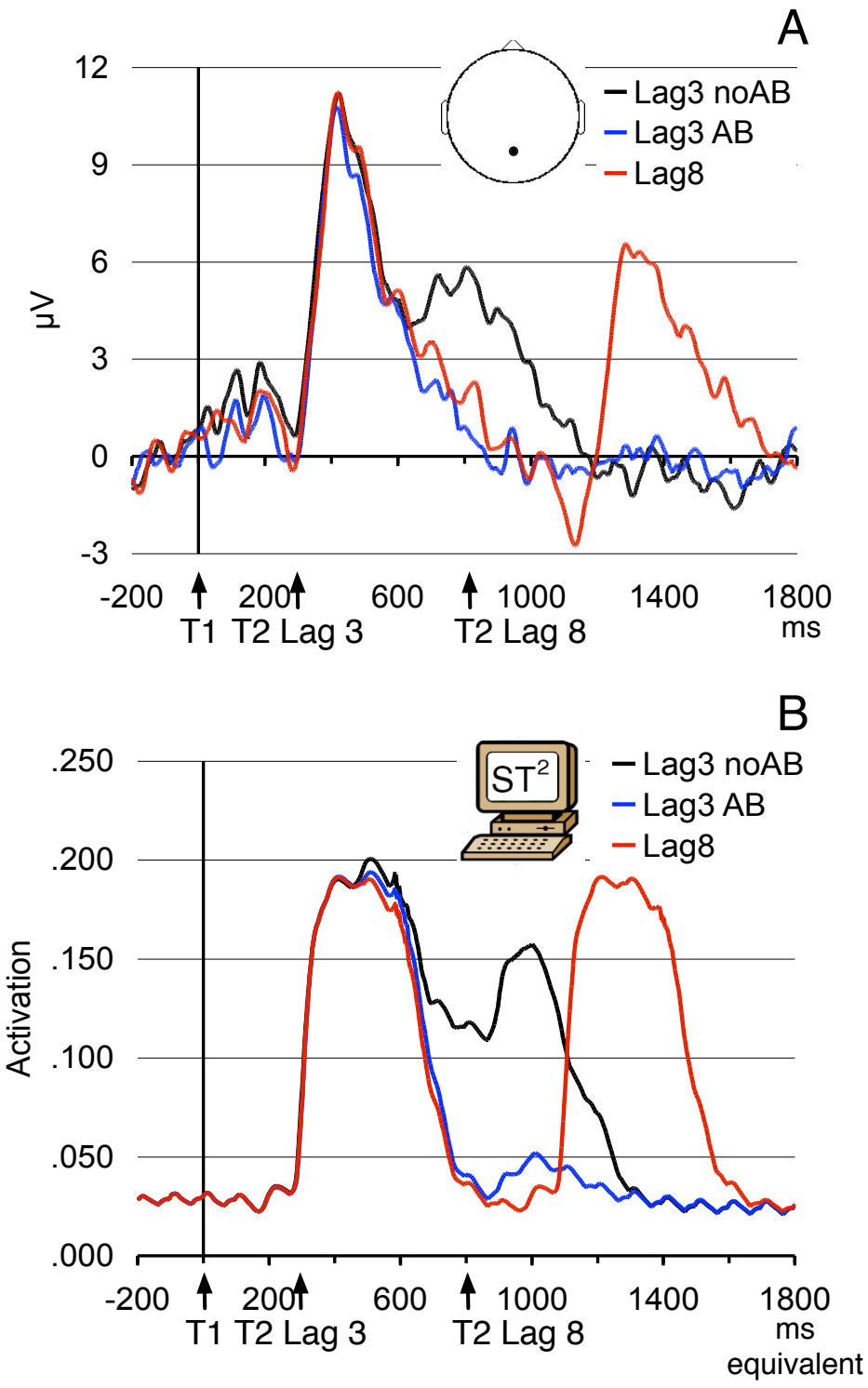


*Figure 38* Panel A: hERP from Pz for 1) the lag 3 noAB condition (T1 and T2 correctly reported), 2) the lag 3 AB condition (T1 accurately identified but T2 not correctly reported), 3) the lag 8 condition (T1 and T2 correctly reported). Positive is plotted upwards. Panel B: $ST^2$'s vERP containing the virtual P3 for 1) the lag 3 noAB condition, 2) the lag 3 AB condition, 3) the lag 8 condition. In both panels 'T1' and 'T2' indicate the presentation of the T1 and T2 respectively, and ERPs are time locked to T1.

does not occur (Figure 38A, see also Kranczioch et al., 2003). The difference in mean amplitude in the 600-1200ms window between the AB and noAB condition is highly significant (Lag 3 AB: 0.7$\mu$V (SEM 0.6) vs. Lag 3 noAB: 3.4$\mu$V (SEM 0.6); $F(1,17) = 24.58$, MSE = 2.6, $p < 0.001$).

Figure 39A suggests the presence of a joint P3 for T1 and T2 if T2 is presented at lag 1. The mean P3 amplitude in the 300-600ms window is significantly larger than the mean amplitude for the same window if T2 is presented at lag 8 (Lag 1: 8.5$\mu$V (SEM 0.5) vs. Lag 8: 7.0$\mu$V (SEM 0.6); $F(1,17) = 11.03$, MSE = 1.77, $p = 0.004$).

### 6.4.3 Virtual ERPs

According to the $ST^2$ model, at lag 3 and lag 8 targets are encoded into working memory in a serial fashion. If T2 is presented at lag 3, the blaster is suppressed by T1's encoding process and T2's tokenisation is delayed. However, since a T2 presented at lag 8 appears after T1 has been encoded into working memory, the T2 can initiate a new encoding process.

As shown in Figure 38B, there is no difference in the mean amplitude of T1's vP3 amplitude, irrespective of whether or not an AB occurs at lag 3 or whether T2 is presented at lag 8 (Lag 3 noAB: 0.18; Lag 3 AB: 0.18; Lag 8: 0.18). There is also no difference in 50% area latency for T1's vP3 component between the lag 3 AB, the lag 3 noAB condition and the lag 8 condition (Lag 3 AB: 470ms equivalent; Lag 3 noAB: 470ms equivalent; Lag 8 470ms equivalent). In line with serial working memory encoding, at lag 3 and lag 8 T2 is presented beyond the time point where it could have an effect on T1's tokenisation.

T2 items that are presented at lag 3 and have relatively low target strength are not encoded into working memory. They show only a small deviation from baseline in the vERP (Figure 38B, T2 vP3 mean amplitude for Lag 3 AB: 0.06), which remains below threshold. T2s that are strong enough to 'outlive' T1's tokenisation, however, re-fire the blaster once T1 encoding has completed. They are consolidated into working memory and show a vP3 component (Figure 38, T2 vP3 mean amplitude for Lag 3 noAB: 0.13).

According to the $ST^2$ model, T1 and T2 are jointly encoded into working memory at lag 1. T2 is presented within the period of T1's blaster enhancement and joins into T1's tokenisation process. Hence, the vERP in Figure 39B contains one joint vP3 component for both T1 and T2 at lag 1. The joint vP3 at lag 1 combines bottom-up activation of two

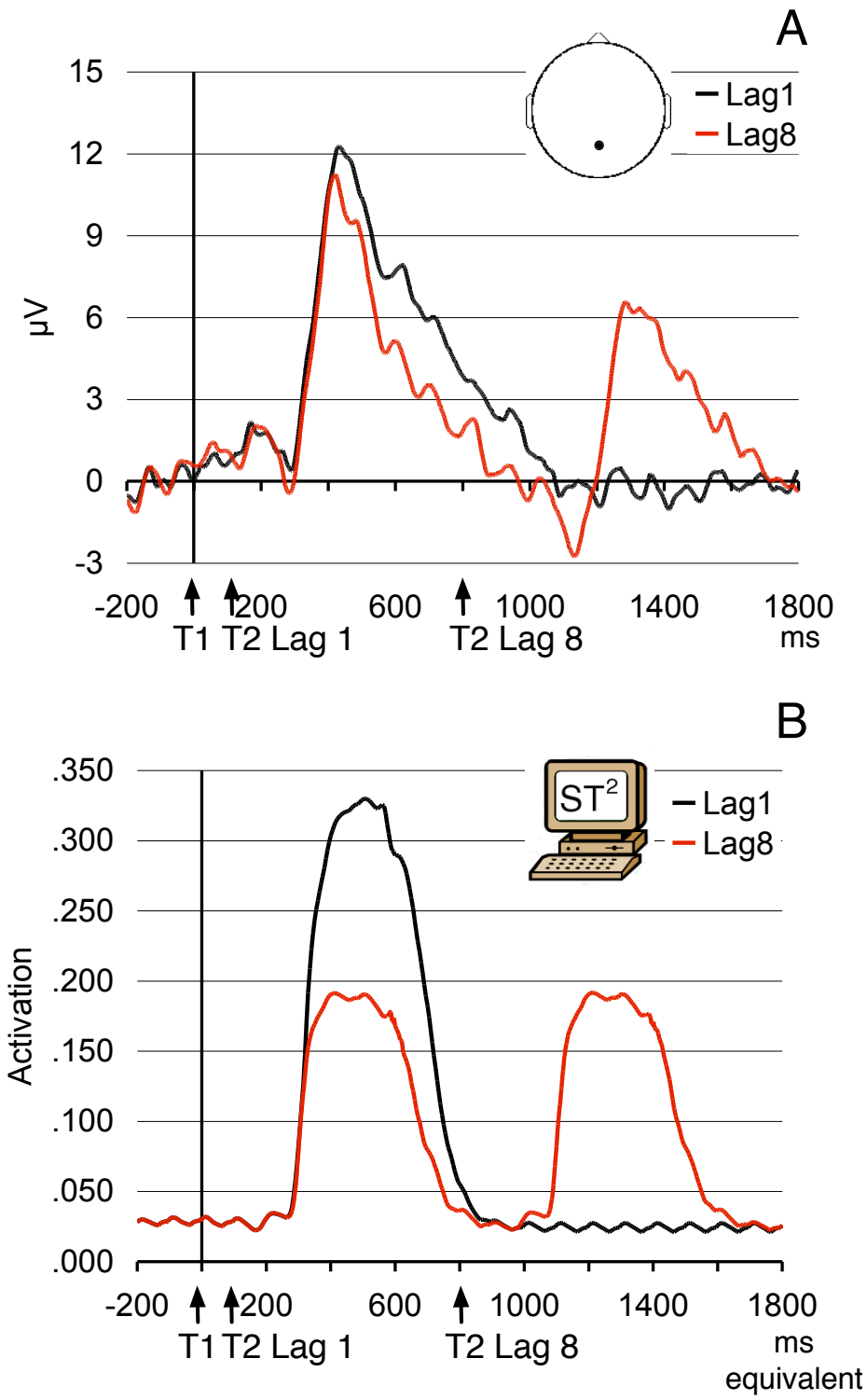


*Figure 39* Panel A: hERP from Pz for 1) the lag 1 condition (T1 and T2 correctly reported), 2) the lag 8 condition (T1 and T2 correctly reported). Positive is plotted upwards. Panel B: $ST^2$'s vERP containing the virtual P3 component for 1) the lag 1 condition, 2) the lag 8 condition. In both panels 'T1' and 'T2' indicate the presentation of the T1 and T2 respectively, and ERPs are time locked to T1.

targets, which is reflected in a larger area under the vP3 curve for the lag 1 vP3 compared to a vP3 for an individual target, i.e. T1's vP3 if T2 is presented at lag 8 (Lag 1: 0.28 vs. Lag 8: 0.17).

## 6.5 Discussion

This chapter addresses two issues central to the evaluation of theories of the AB using electrophysiology. First, we investigate the effect of task difficulty on the P3 component evoked by a target presented in RSVP. Various hypotheses provide conflicting predictions on the relationship between task difficulty and the P3, in that if the target is harder to detect, the amplitude of the P3 should (a) increase (Martens, Elmallah, et al., 2006), (b) remain equal (Shapiro et al., 2006; Kranczioch et al., 2007) or (c) decrease (Kok, 2001). Second, since we do not find a modulation of T1 processing by T2 presented during the AB, our data is in contrast with previously published findings (Shapiro et al., 2006; Martens, Elmallah, et al., 2006; Kranczioch et al., 2007). In the final part of this chapter, we first evaluate our findings (Sections 6.5.1, 6.5.2 and 6.5.3) and then discuss the discrepancy between our data and the previous experimental findings in Section 6.5.4.

### 6.5.1 The meaning of P3 amplitude for targets in RSVP

The results from Experiment 1 provide evidence in favour of the P3 component for targets in RSVP being a correlate of bottom-up target strength. First, certain target letters have significantly higher accuracy scores than others. We use the behavioural data from a previous study (Bowman & Wyble, 2007) to classify target letters as being easy or hard. Our results replicate the previous finding and show a highly significant difference in accuracy between easy and hard letters. This suggests that there are consistent differences in target strengths, which are determined by the identity of each target letter. Such a measure of task difficulty is purely due to intrinsic stimulus characteristics. As target letters are presented at random, observers cannot predict whether a target is going to be easy or hard.

Second, the P3 amplitude is significantly larger for easy compared to hard targets. This finding contradicts theories based on the assumption that P3 amplitude reflects the amount of resource allocated to processing a target in RSVP. According to such theories,

more resource should be required to process harder targets (Martens, Elmallah, et al., 2006). Consequently, although we should find a larger P3 for hard targets, the data from Experiment 1 shows the opposite effect. Alternatively, P3 size might be determined by the amount of resource allocated to the processing of the target, which more or less randomly fluctuates from trial to trial (Shapiro et al., 2006; Kranczioch et al., 2007). However, this hypothesis predicts that a measure of task difficulty due to intrinsic stimulus characteristics (as employed in Experiment 1) should not modulate P3 amplitude, which is in contrast with our results. Hence, based on the results of Experiment 1, we can conclude that if preallocated effort is either random or equal in every trial, as can be assumed due to the randomness of target presentation in RSVP, intrinsic target strength is a main modulator of P3 amplitude.

In neural network terms, target strength might be referred to as bottom-up trace strength. One of the main arguments in the theory underlying the $ST^2$ model is that the working memory encoding process is influenced by the target's bottom-up trace strength. A stronger target will be consolidated into working memory in a more durable manner, which is reflected in a larger vP3 component. Hence, the findings from Experiment 1 validate and support the $ST^2$ model.

### 6.5.2 Working memory encoding is serial during the attentional blink

Both the $ST^2$ model and the resource sharing theory propose that T1 processing affects the consolidation of T2 during the AB, which is supported by behavioural (e.g. Chun & Potter, 1995) and EEG (Vogel et al., 1998) data. In addition to the unidirectional influence of T1 on T2, however, resource sharing also argues that there is mutual interference during the AB, since T1 and T2 compete indirectly through the amount of resource allocated to them. The behavioural and EEG data from Experiment 2, however, do not support this hypothesis. These data suggest that T2 does not influence T1 if presented at lag 3 or lag 8. In addition, there is no effect on T1 processing whether an AB occurs or not.

Our findings support theories that suggest T1 and T2 do not compete for resources during the AB (Olivers, 2007) and are consistent with the hypothesis of serial working memory encoding during the AB (Bowman & Wyble, 2007). If T2 is presented at lag 3, T1 is in the process of being encoded into working memory. During T1's tokenisation process,

the attentional enhancement is suppressed, preventing any interference from T2. Providing T2 has sufficient activation strength, T2's working memory encoding process is delayed until T1 has been consolidated. If T2, however, is too weak, it is lost and an AB occurs.

The data from Experiment 2 is thus in contrast with a key prediction from the resource sharing theory. However, resource sharing - as it stands - lacks a formal interpretation, leaving open the possibility of uncertainty over the exact predictions of the theory. One might thus imagine a modified version of the theory, which would explain the data presented in this chapter, while nevertheless remaining within the 'umbrella' of resource sharing. In that eventuality, however, the resource sharing theory risks becoming 'unfalsifiable'.

### 6.5.3 Interference between T1 and T2 at lag 1

If T1 and T2 are presented in immediate succession (i.e. at lag 1), the serial mechanism of working memory encoding is not enforced. As indicated by the results from Experiment 2, T1 and T2 seem to be encoded into working memory together, thus evoking a single P3 component. This finding is a replication of the MEG results reported in Kessler et al. (2005a), who report a single M300 component for T1 and T2 at lag 1. The increase in swaps at lag 1 provides evidence for joint consolidation during lag 1 sparing, which sometimes leads to a loss of order information for T1 and T2 (Bowman & Wyble, 2007). With respect to the shape of the P3 component at lag 1, neither the human nor the virtual P3 components appear to consist of two individual P3s for T1 and T2 that are offset by 100ms. As the P3 is larger in amplitude but not much broader in time, this suggests a single P3 component (indicating a single enhanced encoding process) for two target items, which is in line with the theory proposed by the $ST^2$ model.

As long as target characteristics are relatively simple (single letters), the joint consolidation has a beneficial effect on T2 accuracy, as exemplified by the lag 1 sparing effect (Bowman & Wyble, 2007). There is a negative effect on T1 accuracy, however, as it is reduced if T2 is presented at lag 1 (see Figure 37 and also Hommel & Akyürek, 2005).

Hence, if there exists some aspect of resource sharing in time, it occurs if targets are presented in immediate succession, as is the case at lag 1. According to the $ST^2$ model, T1 receives an attentional enhancement from the blaster, which lasts for around 150ms. As long as T2 is presented within this period, T2 can join the encoding process and resources

are shared between the two targets.

### 6.5.4 Evaluating previous findings

As previously mentioned, a number of recent articles investigating the AB using EEG and magnetoencephalography (MEG) techniques have argued in favour of resource sharing during the AB. The data from those studies seems to be in contrast with this chapter's findings and predictions from the $ST^2$ model. In the following section, we take a closer look at these previous results. The data presented in each of the articles in question is tested against the following set of criteria, which we believe an EEG/MEG experiment should fulfil in order to provide evidence for resource sharing during the AB:

**P3 as a measure of resource allocation?** Demonstrate that the size of the P3 component evoked by a target in RSVP can be used as a measure of the cognitive resource/effort invested into the detection of that target.

**Resource sharing during the AB?** Resource sharing proposes that if more cognitive resources are allocated to T1, the T2 is more likely to be missed. Accordingly, the P3 component for T1 should be larger for those trials in which an AB occurs compared to when T2 is detected and there is no AB.

#### McArthur, Budd, and Michie (1999)

McArthur et al. (1999) investigate the relationship between T1-related processing (as exemplified by its P3 component) and the AB. Both the P3 component and the AB are 'maximal at about 300 ms' and return to baseline around 600ms following the presentation of T1, thus, it seems that 'the AB and P300 [or P3] follow a similar time course'[2]. Indeed, a significant correlation between the amplitude of six time intervals of the T1 P3 (235-325ms, 328-415ms, 415-505ms, 505-595ms, 595-685ms, 685-775ms; grand averaged across all lags

[2]Note that McArthur et al. (1999) increase the similarity in time course of the P3 component and the AB by shifting the whole AB curve forward in time by 235ms. This is justified by the need to account for 'the propagation delay between probe [the T2] onset and the arrival of the signal [processing related to T1] at the cortex' (McArthur et al., 1999), in order for the T1 to be processed to a level where it could influence the processing of T2.

of T2 presentation) and the depth of the AB[3] at lags 1 - 6 (Figure 2 in McArthur et al. (1999)) emphasises the similarity between the time course of T1's P3 and the AB.

**P3 as a measure of resource allocation?** In McArthur et al. (1999), difficulty is not manipulated on the basis of intrinsic stimulus characteristics (as in Experiment 1 of this chapter) but by making T1 less or more frequent. The authors assume that frequent targets are easy and infrequent targets are hard to perceive. However, the data from Martens, Elmallah, et al. (2006, p. 209) suggests the opposite, i.e. lower average accuracy scores for frequent than infrequent targets, though the results are not significant (p-values of approximately 0.10). Consequently, the relationship between frequency and task difficulty in the AB context is unclear.

Furthermore, due to the very nature of the P3, the less frequent a target is, the more of an 'oddball' it becomes (Kok, 2001). Thus, P3 size is likely to be strongly modulated by frequency/oddball effects, which may not be related to the difficulty of identifying the stimulus, or to the amount of resources allocated to it. With this point in mind, the finding of less frequent targets eliciting a larger P3 (Figure 4 in McArthur et al., 1999) does not per se provide evidence for the P3 component as a measure of resource allocation and does not contradict our results from Experiment 1.

**Resource sharing during the AB?** As T1 P3 data for the Lag 3 noAB condition is not presented in McArthur et al. (1999), this study cannot directly contribute towards the current discussion. However, McArthur et al. (1999) do find a negative correlation between T1 P3 size and depth of the AB ($r = -0.59$, $p = 0.03$, Figure 3 in McArthur et al., 1999), which provides evidence against resource sharing but in favour of a reciprocal relationship during the AB (Bowman et al., 2008).

### Martens, Elmallah, et al. (2006)

This article investigates cueing and frequency effects on the AB. In Experiment 1 of Martens, Elmallah, et al. (2006), T1 difficulty is modulated by making T1 more or less frequent. In Experiment 2 of Martens, Elmallah, et al. (2006), T1 difficulty is manipulated by presenting

[3]The term 'depth of the attention blink' is the opposite of T2 performance, i.e. how strong the AB impairment (and thus how low T2 accuracy) is at that particular point in time.

a cue (the same letter as the T1) above the RSVP stream shortly before the presentation of T1.

**P3 as a measure of resource allocation?** Experiment 1 of Martens, Elmallah, et al. (2006) is a replication of McArthur et al. (1999) in that a notion of task difficulty is modified by making T1 more or less frequent. As stated in the discussion of McArthur et al. (1999), we argue that the relationship between task difficulty and frequency is unclear. What is clear, though, is that frequency alone is a potent factor in determining P3 size (Kok, 2001), which explains a larger P3 for infrequent targets than for frequent targets (Figure 1 in Martens, Elmallah, et al., 2006) without resorting to explanations involving task difficulty or resource allocation.

We believe that the results from Experiment 2 of Martens, Elmallah, et al. (2006) can be explained by the way in which T1 was cued. Although cueing the T1 with the same character makes it easier to detect in behavioural terms, it also makes the T1 less of an oddball, which explains the decrease in P3 amplitude for targets preceded by valid cues compared to invalid cues (Figure 3 in Martens, Elmallah, et al., 2006). Furthermore, invalidly cued T1s also come as more of a 'surprise' to the participant, which increases the amplitude of the P3 component (Donchin, 1981; Kok, 2001). Hence, these results do not per se provide evidence in favour of the P3 being a measure of resource allocation as they are confounded by frequency and expectancy effects influencing P3 amplitude.

**Resource sharing during the AB?** Both experiments presented in Martens, Elmallah, et al. (2006) show T1's P3 to be smaller[4] on those trials in which no AB occurs compared to when T2 is missed and the AB does occur, thus suggesting resource sharing. However, if T1's P3 is mainly modulated by frequency and expectancy effects, as we have suggested in the previous paragraph, the data support a different conclusion. By increasing the frequency of T1 or by validly cueing it, the AB is

[4]Note the effect seems rather weak. In Experiment 1 of Martens, Elmallah, et al. (2006), statistical significance is at p = .085/p = .048 (peak amplitude/400-520ms mean value) when comparing the T1 P3 in the AB to the T1 P3 in the noAB condition. In Experiment 2 of Martens, Elmallah, et al. (2006), significance levels are at p = .050/p = .062 (peak amplitude/432-584ms mean value) when comparing the T1 P3 in the AB to the T1 P3 in the noAB condition.

attenuated (Tables 1 and 2 in Martens, Elmallah, et al. (2006)), which is in line with the reciprocal relationship between T1 strength and the AB (Bowman et al., 2008). Hence, the noAB condition is likely to contain a larger number of frequent T1s (Experiment 1 of Martens, Elmallah, et al. (2006)) and validly cued T1s (Experiment 2 of Martens, Elmallah, et al. (2006)) than the AB condition. Smaller T1 P3s in the noAB compared to the AB condition (Figure 2 and 4 in Martens, Elmallah, et al. (2006)) can be explained by the reduction of T1 P3 amplitude through increased frequency and valid cueing effects. Hence, we argue that the difference in P3 size between the noAB and the AB condition does not per se support resource sharing during the AB.

As it stands, further investigation is needed to provide evidence for resource sharing. Such a study would manipulate task difficulty using intrinsic stimulus characteristics in order to avoid experimental confounds from various factors affecting P3 size.

**Shapiro et al. (2006)**

This study presents M300 (MEG P3 equivalent) data for both T1 and T2 during the AB. Task difficulty is not manipulated and hence cannot be discussed.

**Resource sharing during the AB?** The difference in T1 M300 amplitude between the AB and noAB condition at lag 2 is not significant ($p > 0.05$), hence, on this measure, the data cannot provide evidence for resource sharing. However, the authors do find that T1 M300 amplitude is reduced if T2 is presented inside compared to outside the AB window, which suggests that T2 is able to influence T1 processing during the AB. Such a finding is in contrast with the $ST^2$ model's proposal of serial working memory encoding during the blink. A potential explanation for the finding might be the experimental setup of the study. There is evidence for interference between targets at lag 1, so a T2 presented at lag 2 might be presented close enough to influence T1 processing. Other studies (i.e. Experiment 2 and also Martens, Munneke, et al. (2006)), which use lag 3 as the AB condition, do not find a modulation of T1's P3, hence the evidence is inconclusive.

Shapiro et al. (2006) report a positive correlation over subjects between the size of a subject's T1 M300 and the 'strength' of their AB impairment. They argue that this is evidence for resource sharing, as it indicates that if a subject is able to allocate less resource to T1 (exemplified by a smaller T1 M300) they are able to reduce their AB deficit. However, such a positive correlation between T1 P3 size and depth of the AB was not found in other previously published studies (Martens, Elmallah, et al., 2006; McArthur et al., 1999).

Furthermore, we believe there might be an additional confound. What if certain participants always have smaller M300 components (for both T1 and T2) than other participants? If, as reported for blinkers and non-blinkers (Martens, Munneke, et al., 2006), these participants are also worse at the behavioural task (i.e. have a stronger AB), this would produce the positive correlation observed in Shapiro et al. (2006) emphasising individual differences in the behavioural and MEG data. It requires a study showing a significant positive correlation between T1 M300 (or P3) size and the depth of the AB within each subject, for instance across experimental blocks, to prove resource sharing.

**Kranczioch et al. (2007)**

Kranczioch et al. (2007)'s EEG study of the AB presents data containing the P3 component for T1 and T2. As task difficulty is not manipulated, this issue is not discussed.

**Resource sharing during the AB?** Kranczioch et al. (2007) report a 'significant interaction of the factors *T2 performance* and *time window* [levels T1-P3 window and T2-P3 window] ($F(1,14) = 5.25$, $p = 0.038$)' when T2 is presented at lag 2, i.e. during the AB (see Figure 2B in Kranczioch et al. (2007)). They conclude that 'the T1-related P3 process is larger for trials in which T2 is missed, whilst the T2-related P3 process is smaller in these trials' and that there is resource sharing during the AB.

We argue, however, that the significant interaction does not necessarily provide evidence for resource sharing. The factor *time window* consists of two levels, namely 'T1-P3' and 'T2-P3', whereas the factor *T2 performance* consists of the levels 'T2 seen' and 'T2 missed'. Although the significant interaction indicates a relationship

between T2 performance and P3 time window, such an analysis is not necessarily evidence for a modulation of the 'T1-P3' by the AB.

We illustrate this by performing an equivalent statistical analysis on our data from Experiment 2. A time window ('T1-P3' & 'T2-P3') by T2 accuracy ('T2 seen' & 'T2 missed') interaction analysis on our data is also significant ($F(1,17) = 7.72$, MSE = 3.5, $p = 0.0129$). Two separate paired tests, however, indicate that the interaction is due to a highly significant relationship between T2 accuracy and 'T2-P3' ($F(1,17) = 24.58$, MSE = 2.6, $p < .001$) whereas a comparison of 'T1-P3' and T2 accuracy is not significant ($F(1,17) = 1.91$, MSE = 2.7, $p = 0.185$). Hence, without a paired test between 'T1-P3' and T2 accuracy, the data from Kranczioch et al. (2007) does not necessarily provide evidence for resource sharing.

**Martens, Munneke, et al. (2006)**

Martens, Munneke, et al. (2006) is not directly related to the current discussion as it is primarily concerned with the difference in EEG signatures between so-called 'blinkers' and 'non-blinkers'. Martens, Munneke, et al. (2006) do, however, make an interesting observation concerning T1 P3 latency, which is relevant to the resource sharing discussion.

**Resource sharing during the AB?** Martens, Munneke, et al. (2006) report delayed T1 consolidation if T2 is presented at lag 3 compared to lag 8. This finding suggests that T2 can have some influence on T1 if presented at lag 3, which is intriguing and indeed troublesome for the $ST^2$ model. The reported delay in T1 P3 latency for T2 inside compared to outside the AB, however, resulted from peak latency analysis (lag 3: 495ms, lag 8: 427ms, $t(10) = 2.275$, $p = .046$; S. Martens, personal communication, January 2007). Luck (2005) suggests that if ERP components overlay in time, as is the case during the AB, a 50% area latency analysis (Luck & Hillyard, 1990) can yield more reliable results. Since the present study and others (Shapiro et al., 2006; Martens, Elmallah, et al., 2006; Kranczioch et al., 2007) do not find a delay in T1 consolidation if T2 is presented at lag 3 compared to lag 8, the evidence in favour of delayed T1 consolidation during the AB is inconclusive.

## 6.6 Conclusion

In this chapter, we use the data from Experiment 1 and 2 to address issues fundamental to the evaluation of current theories of temporal attention and the AB. We use the $ST^2$ model to generate vERP traces, which we compare to the hERPs. In addition to validating the dynamics of the computational model, the vERPs are used to make predictions from the theory underlying the $ST^2$ model.

The EEG results presented in Section 6.3 suggest that, at least for targets in RSVP, the P3 component is modulated mainly by target strength and provides only a limited measure of the amount of resource allocated to the task. Thus, EEG/MEG experiments that were taken in support of the resource sharing theory, which assumed P3 size to be a measure of cognitive resource allocated, might have to be reinterpreted.

In Section 6.4 we present EEG results, which suggest that if two targets are presented in immediate succession and within a very short period of time (<150ms), they can be encoded into working memory together. However, during the AB, our data suggest that the encoding of the first target into working memory influences the consolidation of subsequent targets, but this interference is not mutual. Thus, 'resource sharing in time' seems to be limited to short time spans (<150ms) and cannot be extended to the duration of the AB.

To recapitulate the issue of dividing an attentional resource amongst multiple tasks, we can conclude that although such a mechanism seems to exist in the spatial domain (Cavanagh & Alvarez, 2005), resource sharing in temporal attention is severely limited. When orienting in space, the system seems to be able to dynamically adapt its behaviour to achieve an effective trade-off between monitoring the visual field and looking at individual items in detail. In time, however, such dynamic adaptation is restricted to very short time periods (i.e. lag 1) where it is constrained by the length of an attentional episode. Thus, as suggested by the $ST^2$ model, the AB is an observable side effect of this strategy, which enforces a notion of serial order and ensures that perception of stimuli in time is unambiguous.

# Chapter 7

# Temporal variation in target processing during the attentional blink

In this chapter, we investigate the hypothesis that there is increased temporal variance in the deployment of attention and subsequent working memory encoding during the AB. Such a theoretical argument is inherent to the $ST^2$ model and supported by behavioural research on the AB. We compare the EEG patterns evoked by targets inside the AB to those evoked by targets outside the AB and analyse the N2pc and P3 component of the ERP, as these have been associated with attentional selection and working memory encoding, respectively. Using both qualitative and quantitative techniques, we analyse the trial-to-trial variance in the temporal profile of the EEG data underlying the grand average ERP in order to evaluate the hypothesis proposed by the $ST^2$ model. In addition to the analysis of the experimental results, we generate virtual ERPs in each of the conditions. As virtual ERPs are also generated at the single trial level (by means of virtual ERPimages), we can make detailed predictions for the human data and also validate the neural dynamics of the model.

## 7.1 Introduction

During perception of the world humans are constantly faced with an abundance of sensory information. The eyes perform the initial processing of incoming information in the visual domain. As this sensory information feeds through the various layers of visual cortex, it is progressively integrated to gradually generalise over spatial information and level of visual detail (Hochstein & Ahissar, 2002). Whereas early visual areas extract primitive shapes and forms, brain areas situated higher in the visual processing pathway can detect more complex objects. Bottom-up input feeding through this feedforward hierarchical pathway is constantly monitored for salience. The detection of salience triggers an attentional response, which is thought to proceed back down the hierarchy to amplify the neural representation of those salient items (Hochstein & Ahissar, 2002). In other words, if something is deemed to be task-salient, this interrupts the system's ongoing scanning of bottom-up input. As suggested by Nieuwenhuis, Gilzenrat, et al. (2005), the visual system switches from its *tonic* mode of operation, and goes into a focused (or *phasic*) mode of operation. The burst of attention then allows the neural representation of the salient item to rise above surrounding stimuli in terms of neural activation patterns.

This notion of the visual system, which initially generates a coarse-grained representation of the environment and then focuses on specific salient details ('vision at a glance vs. vision with scrutiny'; Hochstein & Ahissar, 2002), is supported by evidence from neurophysiological (e.g. Bar et al., 2006) and behavioural studies. For instance, various behavioural studies have found that participants were able to detect words before individual letters (Johnston & McClelland, 1974), scenes before individual objects (Biederman, Rabinowitz, Glass, & Stacy, 1974) and forests before trees (Navon, 1977).

### 7.1.1 Transient attention and the $ST^2$ model

In the context of temporal visual processing, such a transient attentional enhancement is thought to ensure that fleeting stimuli, which are relevant for the task at hand, have a good chance of reaching conscious perception (Bowman & Wyble, 2007). As discussed in Section 3.3, the $ST^2$ model suggests that salient stimuli trigger a transient attentional enhancement (the blaster), which amplifies the type representation of the stimulus. The

blaster often provides the stimulus with sufficient activation to be bound to a token. During this tokenisation, however, the blaster is suppressed, which prevents the working memory encoding process from being corrupted by the intrusion of other salient items. The phasic mode of attentional selection through the blaster thus creates episodes of working memory encoding (Wyble et al., 2009).

Such circumstances occur during the AB, where items are presented in rapid succession and the second of two targets (T1 and T2) is often missed. Detection performance is excellent, however, if T1 and T2 are presented in immediate succession (lag 1 sparing; see Section 2.3.2). In Chapter 6, we presented evidence suggesting that at lag 1, targets are encoded in a single attentional episode. At lag 3 and due to the serial nature of working memory encoding, however, T2 misses the episode initiated by T1, which results in the AB. Accordingly, the AB can be seen as an artifact of the phasic mode of transient attention and seems to provide insights into the length of an attentional episode.

### 7.1.2 Increased temporal variance in target processing during the AB?

Behavioural studies have suggested that there is increased temporal variance in the deployment of attention and consequent encoding of items into working memory during the AB. Specifically, it was found that during the AB observers often perceive neighbouring items in the RSVP stream instead of T2 (Popple & Levi, 2007). In fact, in the analysis of Popple and Levi (2007), if a trial is categorised as *correct* if participants report the identity of either T2 itself or one of the items presented before or after T2 ($\pm 3$ positions in the RSVP stream), the AB disappears. In addition, subjects often make binding errors when a T2 presented during the AB consists of multiple features (Chun, 1997a), so-called *illusory conjunctions* (Botella, Barriopedro, & Suero, 2001).

These behavioural results support the $ST^2$ model, which proposes that the increased temporal variance of attentional deployment during the AB is due to the episodic nature of the visual system. From this perspective, the AB is an artifact of the visual system attempting to assign unique tokens to targets. As discussed in Section 3.3.3, the blaster is suppressed during T1's tokenisation. The period of blaster unavailability varies from trial to trial depending on how long it takes to tokenise T1, which in turn depends on the strength of T1. If T2 is presented within 600ms of T1 (as is the case during the AB)

and the T2 is too weak in terms of bottom-up trace strength to be tokenised without an enhancement from the blaster, T2 will be missed on trials in which T1's tokenisation takes too long. However, on some trials, T2 is strong enough to 'outlive' this period of attention being unavailable. In this case, the temporal dynamics of T2's encoding process will vary depending on how long it has taken the system to tokenise T1. Hence, the $ST^2$ model proposes that for targets that are correctly reported during the AB (at lag 3), there should be increased trial-by-trial variance in the deployment of attention and subsequent working memory encoding, when compared to targets outside the AB (i.e. targets presented outside the 'blink-window' at lag 8).

### 7.1.3 Overview

In this chapter, we investigate this hypothesis by comparing the EEG signatures evoked by targets presented inside the AB to those evoked by targets outside the AB. We analyse whether there is increased temporal variance in the attentional response and subsequent working memory encoding of targets presented inside the AB. For this purpose we are interested in the N2pc component, which is considered to be an EEG correlate of selective attention to salient information, and the P3 component, which is commonly associated with encoding items into working memory (see also Section 2.1.1).

Previous studies have investigated how the grand average P3 and N2pc components for targets presented inside the AB differ compared to the P3 and N2pc components evoked by targets presented outside the AB. It has repeatedly been found that the P3 is both attenuated in terms of component amplitude (Vogel et al., 1998; Kranczioch et al., 2003; Sessa, Luria, Verleger, & Dell'Acqua, 2006) and also delayed in terms of component latency (Vogel & Luck, 2002; Sessa et al., 2006) for targets inside compared to outside the AB. Similarly, Jolicoeur, Sessa, Dell'Acqua, and Robitaille (2006) report a latency delay and amplitude attenuation of the N2pc component amplitude for targets inside the AB when compared to targets outside the AB.

However, the grand average N2pc and P3 ERP components cannot directly elucidate the investigation of our hypothesis, as the averaging process collapses across temporal fluctuations in individual EEG trials. Specifically, given a set of trials that are averaged together, both decreases in amplitude and increases in latency variation will attenuate the

mean amplitude of the grand average ERP and possibly cause it to be broader, i.e. more 'spread out' in time. Hence, any trial-by-trial variation that might have been present in the ERP results presented in the aforementioned studies would indeed 'wash out' in the grand average ERPs. Consequently, in the folllowing we employ visualisation techniques that allow us to investigate the single trial dynamics underlying ERPs (see Section 2.1.1). We believe that, when employed in conjunction with concrete a priori hypotheses, such analyses provide effective means for testing these hypotheses.

In addition to the human EEG data, we generate virtual ERP components, which are hypothesised to be the $ST^2$ model's equivalent of the N2pc and P3 components of the human ERP (see Section 4.5). For each of the experimental conditions, the virtual ERPs are contrasted with the human data. This comparative evaluation allows us to validate the $ST^2$ model and propose explanations for the human ERP effects.

## 7.2 Methods

This chapter is based on behavioural and EEG data from Experiment 2. Please refer to the appendix for a detailed overview of the methods employed in this experiment. Details specific to the analyses presented in this chapter are described in the following sections.

### 7.2.1 EEG analysis

For the analyses of this chapter, the continuous EEG data from each participant are first low-pass filtered at 25Hz and then segmented into trials. This is done by extracting a time window of -500ms to 1500ms around the target onset times for the conditions of interest, namely *seen T2s at lag 8 following a seen T1*, and *seen T2s at lag 3 following a seen T1*. After artifact rejection, the total number of trials in the above conditions of interest are: T2 at lag 8: 938: T2 at lag 3: 853. The segmented EEG data are detrended to remove direct current drift artifacts. For the 'T2 at lag 8' condition the trials are baselined to the -200ms to 0ms window preceding T2 presentation. For T2s presented at lag 3, however, this window coincides with T1's P3 and hence we baseline to the -200 to 0ms window with respect to the presentation of the preceding T1 (i.e., -500 to -300ms with respect to the T2 at lag 3). Activity from the Pz (midline parietal) electrode is used to analyse the P3

component. The N2pc component is calculated as the difference waveform obtained by subtracting out the average ipsilateral waveform from the average contralateral waveform. For a target presented to the left (right) of the fixation cross, the contralateral (ipsilateral) waveform is calculated by averaging across the P8 and O2 electrodes, and the ipsilateral (contralateral) waveform is calculated by averaging across the P7 and O1 electrodes.

As stated in the methods in the appendix, the $ST^2$ model's simulation run of the two-target paradigm in Experiment 2 contains 169 trials. The virtual ERPs for the conditions of interest for this chapter (*seen T2s at lag 8 following a seen T1*, and *seen T2s at lag 3 following a seen T1*) contain 121 and 44 trials, respectively.

### ERPimage analysis

As described in Section 2.1.1, ERPimages visualise EEG activity over time across the individual trials in a condition of interest. The goal of our analysis is to compare the amount of temporal variance underlying the ERP components for targets inside and outside the AB. However, the *temporal profile* of an ERP component cannot be clearly distinguished at the single trial level.

In line with our hypothesis, increased temporal variance underlying an ERP component should result in increased variation in the phase of the single-trial EEG waves underlying the grand average. Consequently, we convert our EEG data to the frequency domain and investigate the phase values of the individual trials at the frequency and time point of the ERP component of interest. We approximate the frequency and profile of the wavelet to be used for the time-frequency decomposition from the grand average profile of the ERP component of interest. The P3 and N2pc components can be approximated as half cycle waves. In terms of frequency, the P3 typically lasts for just over 300ms, which corresponds to an approximated frequency of 1.5Hz. The N2pc, on the other hand, typically lasts for roughly 200ms, so we approximate the frequency of the N2pc to be 2.5Hz. Consequently, we choose a half-cycle wavelet at 1.5Hz for the P3 and a half-cycle wavelet at 2.5Hz for the N2pc analysis. The time-frequency analysis provides us with a matrix of phase values for each time point in each individual trial. To determine a latency of the ERP components, we let the algorithm pick the point at which the amplitude of the grand average ERP hits its maximum for the P3 (and minimum for the N2pc) during the relevant time window.

Then, we extract the phase value at this time point of the peak of the ERP component, which results in an array containing one phase value per trial.

These phase values are then fed into the corresponding EEGlab function (Delorme & Makeig, 2004) to sort the trials of the ERPimages by their phase values. A dashed vertical line in the ERPimage plot indicates the time point at which the phase is extracted, i.e. the peak of the ERP component. To enhance the visualisation of the ERPimage, we then vertically smooth across a moving window of 50 trials.

## 7.3 Results & discussion

The following section contains the EEG results evoked by targets inside and outside the AB. In the final part of the section, we use the $ST^2$ model to generate virtual ERPs and discuss the implications of the comparison with the human ERPs for the theory underlying the $ST^2$ model.

### 7.3.1 Behavioural results

In the following analysis, we compare the EEG correlates of targets presented outside the AB (T2 following a seen T1 at lag 8) to targets presented inside the AB (T2 following a seen T1 at lag 3). To reiterate the behavioural results that were already presented in Chapter 6, human accuracy for targets outside the AB is 74% (SEM 4). A target inside the AB is correctly detected on 54% (SEM 5) of the trials. The difference is significant, $F(1,17) = 11.67$, $MSE = 0.03$, $p = 0.003$.

### 7.3.2 Evidence for increased variance during the AB

In the following, we present ERPimages sorted by the phase values of the individual trials, as described in the methods of this chapter. These allow a qualitative inspection of the amount of variance of the EEG underlying the P3 and N2pc components for targets inside and outside the AB.

However, the qualitative effect needs to be backed up by quantitative evidence to support the hypothesis of increased variance in target processing during the AB. To this end, we perform a statistical analysis of the distribution of phase values calculated at the peak of the

ERP component, as described in the methods section of this chapter. The distribution of phases is modelled as a circular von Mises distribution with a circular mean $\mu$ and a measure of concentration $\kappa$ (Evans, Hastings, & Peacock, 2000). $\kappa$ can be seen as the analogue of the inverse of the variance of the distribution. A larger $\kappa$-value means that a distribution is more centred around the mean and hence has lower variance than a distribution with a smaller $\kappa$-value, which indicates that the distribution is more spread out. The $\kappa$-value for a distribution can thus be used to investigate our hypothesis, in that we expect the phase distribution for targets inside the AB to have a smaller $\kappa$-value than the phase distribution for targets outside the AB. To test the comparisons of $\kappa$-values for statistical significance, we estimate the $\kappa$-values for each subject in each condition using maximum likelihood estimation and then feed the results into a repeated measures ANOVA.

**P3 ERPimage** The ERPimages in Figure 40 depict the single-trial EEG underlying the grand average P3 from the ERP trace shown below each of the ERPimages. As described in the methods, the trials within the ERPimage are sorted by the phase value at the approximated time point and frequency of the P3 component. Visual inspection suggests that the positive (red) activation corresponding to the P3 component is more vertically aligned for targets outside the AB (Figure 40A) than for targets inside the AB (Figure 40B), suggesting that the phases of the EEG trials are relatively similar near the peak time and frequency of the P3 ERP component. Conversely, the fact that the P3-related EEG activity throughout trials is more diagonally sloped for targets inside the AB suggests that there is more variance in the distribution of phase values for these targets when compared to targets outside the AB.

The qualitative finding of increased variance during the AB is supported by the statistical analysis of the $\kappa$-values for the phase distributions for targets inside and outside the AB. Whereas the mean $\kappa$-value across participants for targets outside the AB is 0.96 (SEM 0.09), the mean $\kappa$-value for targets inside the AB is 0.52 (SEM 0.10). The difference in mean $\kappa$-values between the two conditions is significant: $F(1,17) = 15.6$, MSE = 0.12, $p = 0.001$.

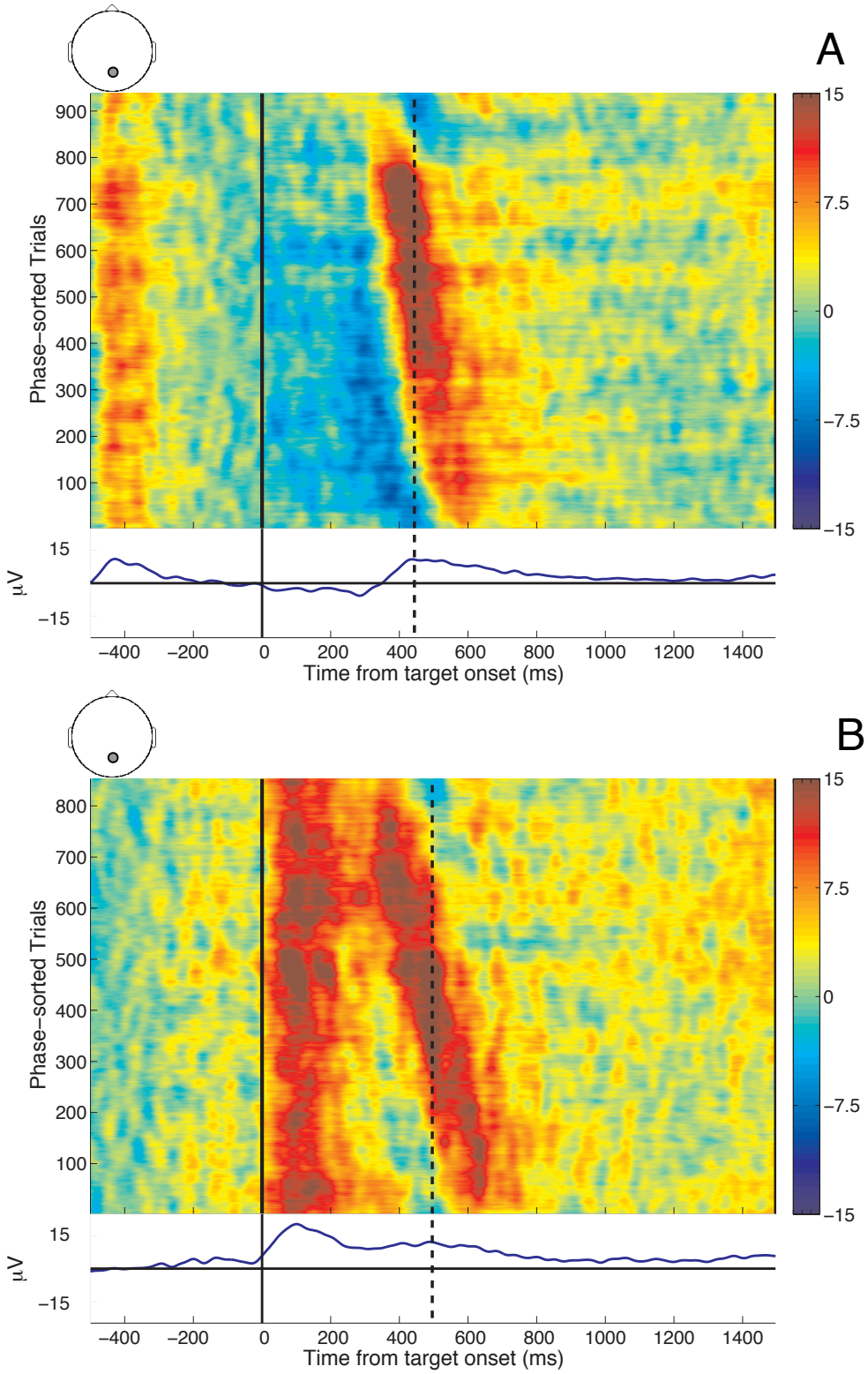


*Figure 40* Panel A: P3 phase-sorted ERPimage for targets outside the AB (T2 at lag 8). Panel B: P3 phase-sorted ERPimage for targets inside the AB (T2 at lag 3). The T2 is presented at time point zero. The dashed line indicates the time point around which the phase sorting is centred, corresponding to the peak of the P3 component.

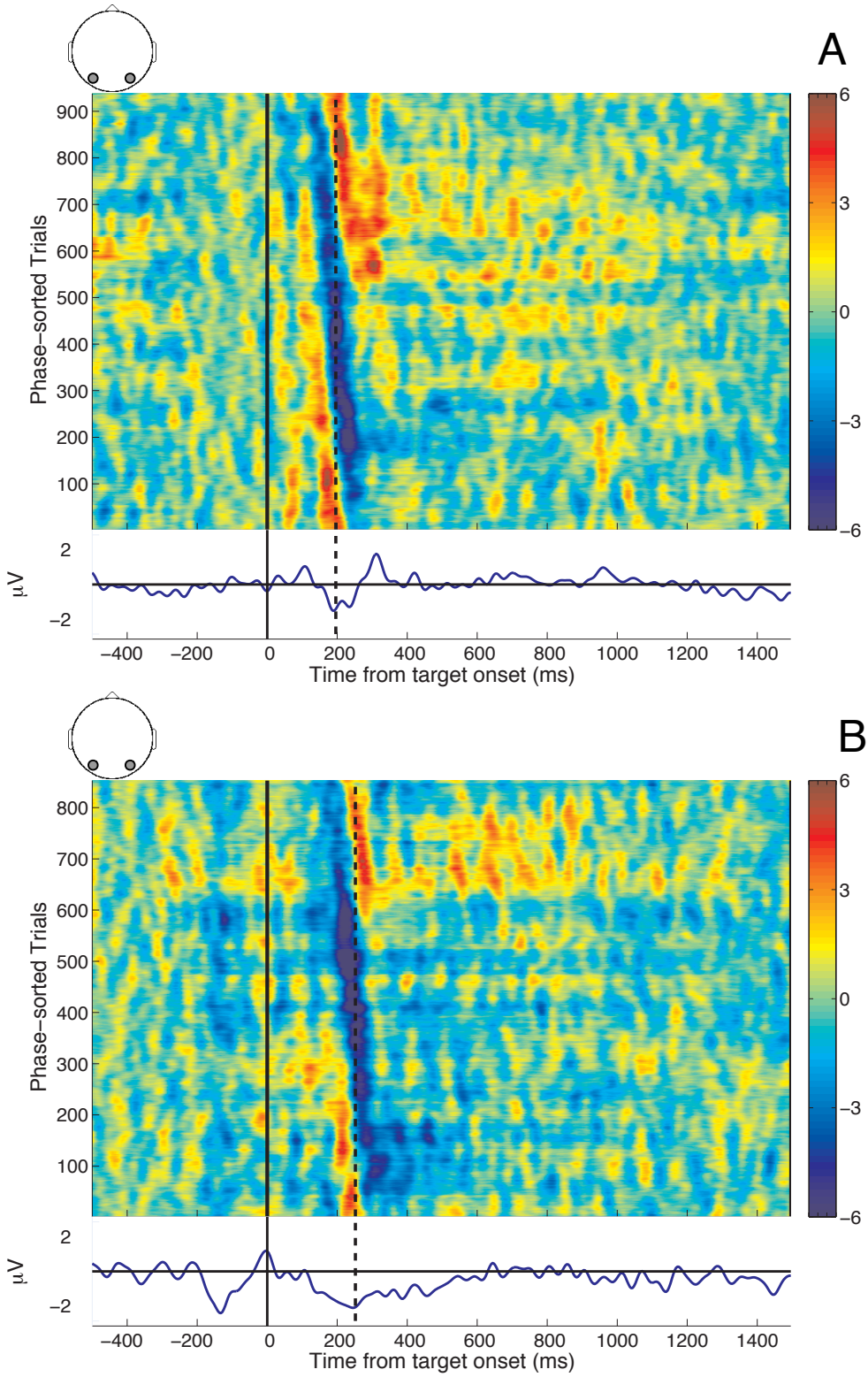


*Figure 41* Panel A: N2pc phase-sorted ERPimage for targets outside the AB (T2 at lag 8). Panel B: N2pc phase-sorted ERPimage for targets inside the AB (T2 at lag 3). The T2 is presented at time point zero. The dashed line indicates the time point around which the phase sorting is centred, corresponding to the peak of the N2pc component.

**N2pc ERPimage** The ERPimages in Figure 41 visualise the single-trial EEG underlying the grand average N2pc shown below it. Although admittedly a weaker effect than for the P3 component, the visual comparison of the ERPimages suggests that across all trials the negative (blue) activation corresponding to the N2pc evoked by targets inside the AB (Figure 41B) shows more of a diagonal slope than the N2pc activation for targets outside the AB (Figure 41A). This indicates that there is more variance in the distribution of phase values for targets outside the AB than inside the AB.

The statistical analysis of the N2pc phase distributions supports the qualitative finding of increased variance for targets inside the AB. For targets outside the AB, the mean $\kappa$-value across participants is 0.50 (SEM 0.04) and the mean $\kappa$-value for targets inside the AB is 0.33 (SEM 0.05). The difference in mean $\kappa$-values between the two conditions is significant: $F(1,17) = 9.60$, MSE = 0.03, $p = 0.007$.

### 7.3.3 Verifying our analysis of phase distributions

The results presented in the previous section thus seem to support our hypothesis of increased variance in target processing during the AB. However, there is a potential confound, which would invalidate the results from our phase distribution analysis.

The grand average P3 and N2pc components for targets inside and outside the AB show differences in ERP component amplitude. Hence, the condition with the lower amplitude of the ERP component might suffer from a too small signal-to-noise ratio, in terms of how the EEG activity from the individual trials contributes to the grand average ERP. If the signal-to-noise ratio was indeed below a certain critical level, the time-frequency analysis might not be able to correctly detect the phase of the ERP component at the individual trial level but instead measure the phase of random EEG noise. This would result in a smaller $\kappa$-value for that phase distribution, which would simply be due to the decrease in the amplitude of the ERP component and would not be due to a more 'spread out' phase distribution. Consequently, we have to verify that our observed differences in $\kappa$-values for targets inside and outside the AB are indeed due to differences in the profile of the phase distributions and not due to differences in ERP component amplitude.

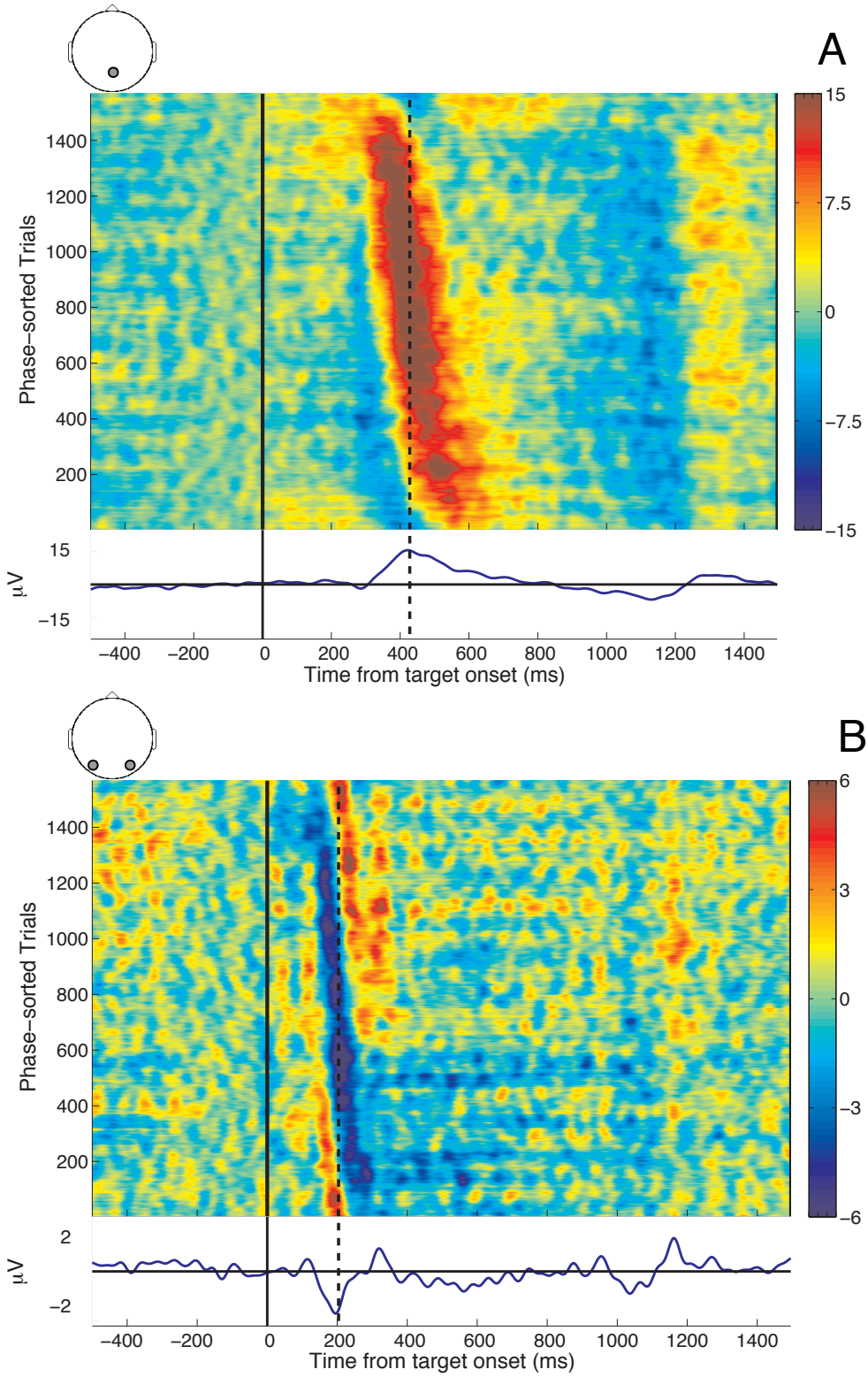


*Figure 42* Panel A: P3 phase-sorted ERPimage for T1 presented at lag 8. Panel B: N2pc phase-sorted ERPimage for T1 presented at lag 8. The T1 is presented at time point zero. The dashed line indicates the time point around which the phase sorting is centred, corresponding to the peak of the P3 and N2pc component respectively.

**1. Differences in ERP component amplitude alone do not explain the reduction in $\kappa$-values**

We have to show that smaller ERP component amplitude alone does not explain the reduction in $\kappa$-values. This can be done by comparing T2 at lag 8, i.e. the target outside the AB condition, to the EEG profile of a T1 preceding T2 presented at lag 8.

From the grand average ERP, we know that, at lag 8, T2s P3 and N2pc components have smaller amplitude than those of the preceding T1. For the P3 component, the difference is significant (T1 lag 8 mean 6.1$\mu$V (SEM 0.5) vs. T2 lag 8 mean 3.9$\mu$V (SEM 0.9); $F(1,17) = 8.21$, MSE = 6.3, $p = 0.011$). Although there is a rather large difference in amplitude for the N2pc component, the ANOVA fails to reach significance (T1 lag 8 mean -0.8$\mu$V (SEM 0.2) vs. T2 lag 8 mean -0.5$\mu$V (SEM 0.3); $F(1,17) = 1.64$, MSE = 0.5, $p = 0.218$).

The ERPimage for the P3 and N2pc evoked by T1 with T2 at lag 8 (shown in Figure 42), however, looks remarkably similar to the EEG profile evoked by the T2 presented at lag 8 (see Figure 40A and Figure 41A). In both plots the activation related to the ERP component is strongly aligned along the vertical dashed line, which suggests little variance in the phase distribution. This qualitative observation is supported by the statistical analysis of $\kappa$-values. For the P3 component, the T1 at lag 8 has a mean $\kappa$-value of 1.01 (SEM 0.09) and for T2 at lag 8 the mean $\kappa$-value is 0.96 (SEM 0.09) across all participants. The ANOVA confirms that this small difference is not significant: $F(1,17) = 0.39$, MSE = 0.04, $p = 0.540$). Similarly, there is little difference in the $\kappa$-values for the N2pc component (T1 lag 8 mean 0.46 (SEM 0.04) vs. T2 lag 8 mean 0.50 (SEM 0.04); $F(1,17) = 0.28$, MSE = 0.04, $p = 0.606$).

To summarise, at lag 8 the P3 and N2pc components have smaller amplitudes for T2 than for T1. Nevertheless, the ERPimages and the statistical analysis of the mean $\kappa$-values suggest that there is no reliable difference in the profile of the phase distributions for these two conditions. Hence, this suggests that ERP component amplitude reduction alone cannot explain the observed reduction in $\kappa$-values when comparing the phase distributions for targets inside and outside the AB.

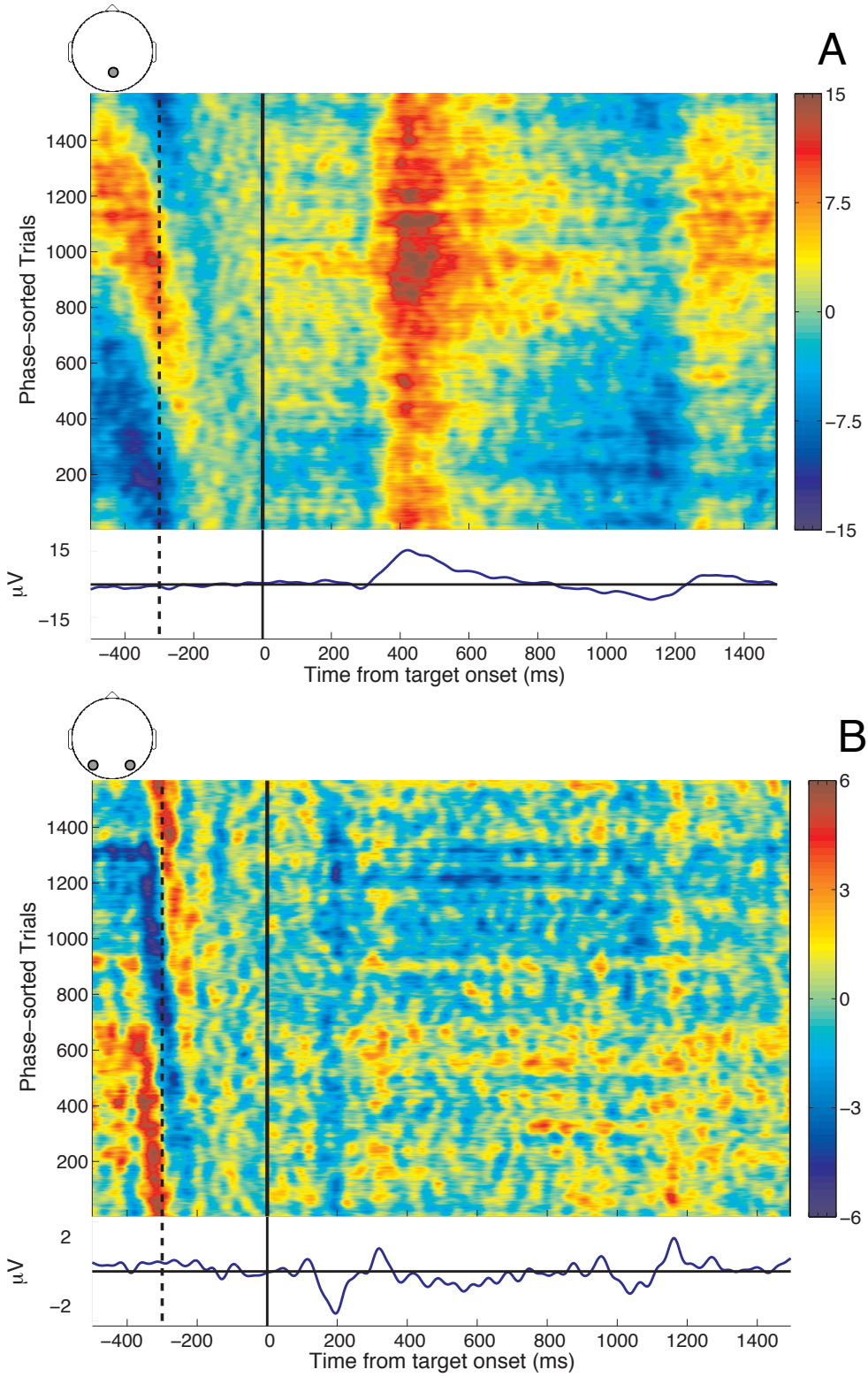


*Figure 43* Panel A: P3 phase-sorted ERPimage for the baseline condition (T1 at lag 8 condition but sorted around the -300ms time point with respect to T1 onset). Panel B: N2pc phase-sorted ERPimage for the baseline condition. The T1 is presented at time point zero. The dashed line indicates the time point around which the phase sorting is centred.

**2. Targets inside the AB have significantly larger $\kappa$-values than baseline**

A second potential confound is that the EEG profile for targets inside the AB (T2 at lag 3) condition might not be distinguishable from background activity, which, if this was the case, would also explain a lower $\kappa$-value for this condition when compared to the phase distribution of targets outside the AB. However, our analysis can be validated against this confound by showing that the $\kappa$-value of the phase distribution for targets inside the AB is significantly higher than the $\kappa$-value of the distribution of phases during baseline EEG activity.

To this end, we compare the 'targets inside the AB' condition (Figure 40B and Figure 41B) to the phase-sorted ERPimage centred around a time point prior to T1 presentation (i.e., -300ms with respect to T1 presentation, see Figure 43). At -300ms with respect to T1 presentation, the EEG shows no consistent stimulus-related activity, which is reflected as zero activation in the grand average ERP. This condition can thus act as a baseline condition where the phases of the EEG trials should be approximately uniformly distributed. And, as a matter of fact, the activation along the vertical dashed line in Figure 43 is indeed more or less uniformly distributed.

If we now compare the $\kappa$-values from the underlying phase distributions, we find a significant difference between targets inside the AB and the baseline condition. For the P3 component, targets inside the AB have a mean $\kappa$-value of 0.52 (SEM 0.10) and a mean $\kappa$-value of 0.19 (SEM 0.03) for the baseline condition across all participants ($F(1,17) = 12.52$, $MSE = 0.08$, $p = 0.003$). Similarly, the mean $\kappa$-value for the phase distribution underlying the N2pc component is significantly smaller for the baseline condition (mean 0.20, SEM 0.04) than for targets inside the AB (mean 0.33, SEM 0.05); $F(1,17) = 5.01$, $MSE = 0.03$, $p = 0.039$.

Consequently, the phase distribution of targets inside the AB has a significantly larger $\kappa$-value than that for baseline EEG activity. This suggests that although the distribution of phases for targets inside the AB is less concentrated than for targets outside the AB, it is still significantly more concentrated than the phase distribution observable in background EEG activity.

### 7.3.4 Virtual ERPs from the $ST^2$ model

As mentioned in the introduction, the hypothesis of increased temporal variation of target processing during the AB is inherent to the $ST^2$ model. In the following section, we generate virtual N2pc and virtual P3 components for targets inside and outside the AB. This approach allows us to validate the internal dynamics of the $ST^2$ model and provides us with theoretical explanations for the human EEG effects.

#### Simulated behavioural accuracy

The simulated behavioural accuracy from the $ST^2$ model is 85% for targets outside the AB (T2 at lag 8) and 31% for targets inside the AB (T2 at lag 3). The $ST^2$ model thus qualitatively replicates the human behavioural data.

#### Virtual ERPimages

As with the analysis of the human ERPs, the average virtual ERP waveforms (as plotted in the previous chapters) are 'blind' to underlying trial-by-trial fluctuations. Hence, we generate virtual ERPimages (see Section 4.3.5), which illustrate the activation profiles of individual trials during a simulation run of the $ST^2$ model. Analogous to the phase sorting of the human ERPimages, the virtual ERPimages are sorted by the peak latency of the virtual ERP component of interest (indicated by the solid black line in Figures 44 and 45) in each trial.

**Virtual P3 ERPimage** The virtual P3 ERPimage shows how the activation for correctly identified targets outside the AB (Figure 44A) is more aligned in time compared to that for correctly identified targets inside the AB (Figure 44B). We can quantify this difference by comparing the mean and standard deviation of the distribution of peak latencies of the virtual P3 across all simulation trials. We find that for targets outside the AB (T2 at lag 8), the mean and standard deviation are 489.0ms (equivalent) and 82.5ms (equivalent), respectively. In comparison, for targets inside the AB (T2 at lag 3), the mean latency is 611.7ms (equivalent) with a standard deviation of 123.4ms (equivalent). The relative increase in the standard deviation reflects increased temporal variation of the virtual P3

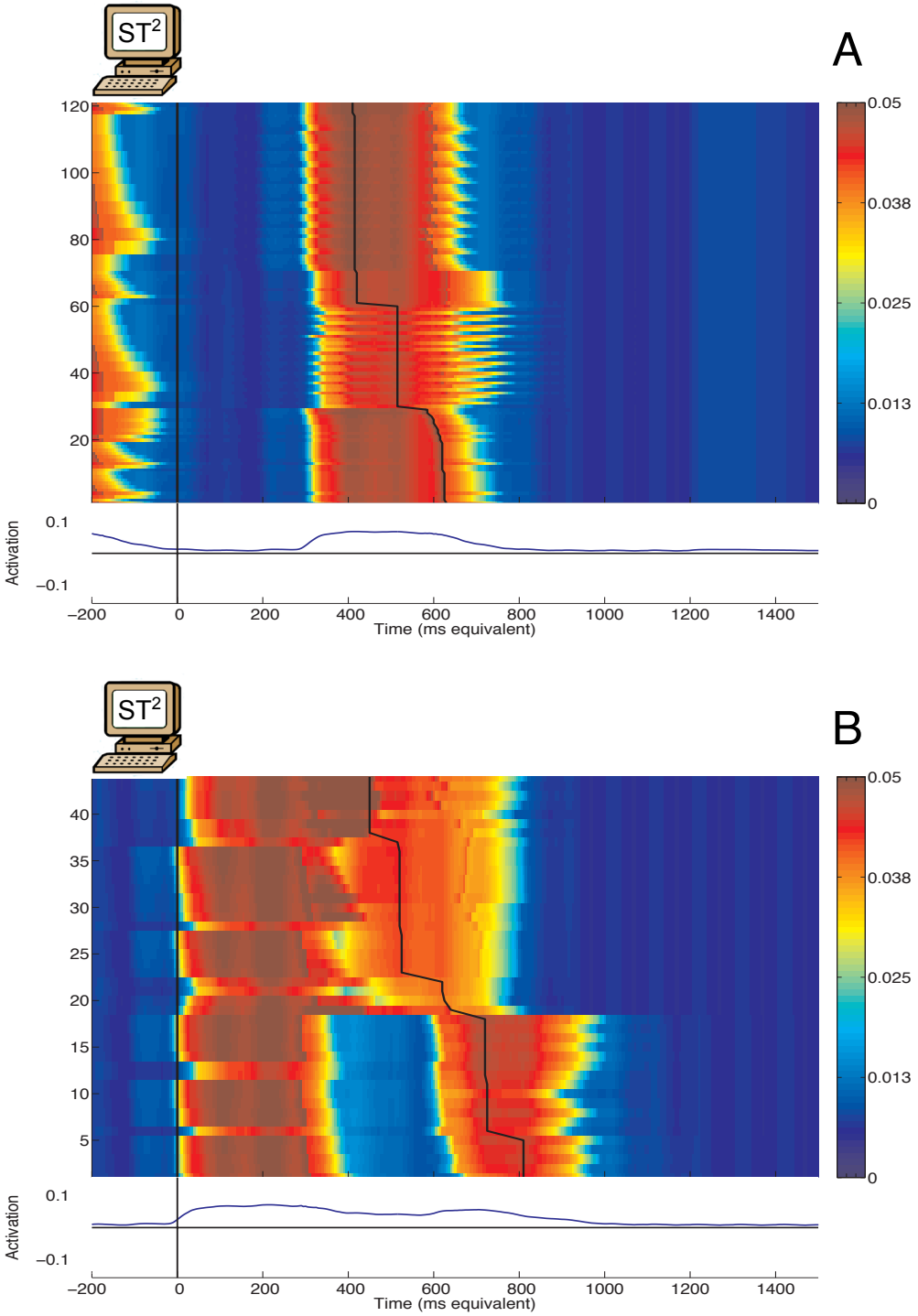


*Figure 44* Panel A: P3 virtual ERPimage for correctly identified targets outside the AB (T2 at lag 8). Panel B: P3 virtual ERPimage for correctly identified targets inside the AB (T2 at lag 3). The target is presented at time point zero. The solid line indicates the peak of the virtual P3 component in that trial.

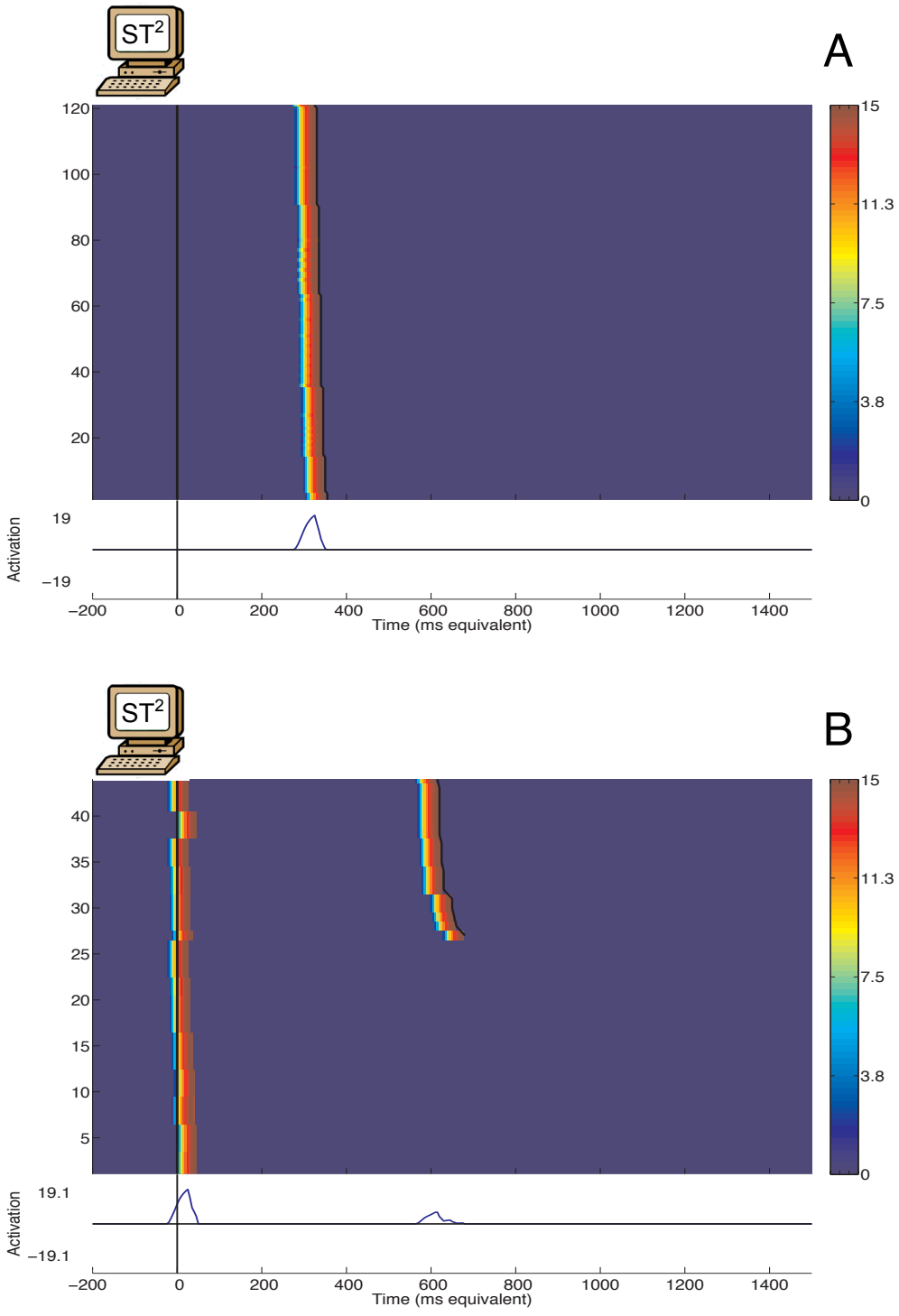


*Figure 45* Panel A: N2pc virtual ERPimage for correctly identified targets outside the AB (T2 at lag 8. Panel B: N2pc virtual ERPimage for correctly identified targets inside the AB (T2 at lag 3). The target is presented at time point zero. The solid line indicates the peak of the virtual N2pc component in that trial.

component for targets inside the AB. Hence, there is a qualitative correspondence between the virtual ERPimages and their human equivalents from Figure 40.

**Virtual N2pc ERPimage** Similarly, the virtual ERPimages of the virtual N2pc show temporally coherent activation for correctly identified targets outside the AB (Figure 45A), whereas for correctly identified targets inside the AB (Figure 45B) it is more 'jittered' in time. The reader will note that in roughly 50% of the trials in Figure 45B, the T2 has enough bottom-up trace strength to be correctly identified inside the AB without any enhancement from the blaster. Hence, these trials (the lower part of Figure 45B) show no virtual N2pc activity related to the T2. As in the analysis of the virtual P3, we compare the mean and standard deviation of the distribution of peak latencies of the virtual N2pc component. We find that for targets outside the AB (T2 at lag 8), the mean and standard deviation are 338.4ms (equivalent) and 7.0ms (equivalent), respectively. In comparison, for targets inside the AB (T2 at lag 3), the mean latency is 633.1ms (equivalent) with a standard deviation of 17.8ms (equivalent). Again, this is a qualitative replication of the pattern of effects in the human N2pc ERPimage (Figure 41).

### Discussion

The virtual ERPimages provide a means of visualising the theoretical framework of the $ST^2$ model at a fine-grained level of detail. Using this methodology of comparing model and human data at the level of single trials, we can show that, in line with the $ST^2$ model's hypothesis, the neural activation traces of nodes corresponding to the deployment of attention and consequent working memory encoding show increased temporal variance for targets inside the AB compared to targets outside the AB.

As previously stated, the blaster is always available for targets presented outside the AB. The blaster fires as soon as an item is classified as a target and provides a burst of activation to aid the tokenisation of the target. The virtual N2pc component, which consists of neural activation from the blaster, is thus temporally well aligned for a target outside the AB (a T2 at lag 8), as the blaster fires at more or less the same time in every trial. Due to the availability of the blaster for targets outside the AB, there is also little variation in the temporal profile of tokenisation and the virtual P3 occurs at approximately

the same latency across trials. In most circumstances, the target's type representation is bound to a working memory token, which results in this target being correctly reported by the $ST^2$ model. Hence, the model produces high simulated behavioural accuracy at detecting targets outside the blink.

However, the processing of a target presented during the AB (a T2 at lag 3), is complicated by the occurrence of the preceding T1. The blaster is suppressed while T1 is encoded into working memory. Thus, the time point at which attention becomes available for T2 depends on the duration of T1's tokenisation process. This in turn is determined by T1's strength, which varies from trial to trial. There is significantly more trial-to-trial variability in the time point at which the blaster eventually fires for T2, as this depends both on how long it takes to encode T1 as well as T2's own strength. Note that on a number of trials, the T2 presented inside the AB has sufficient bottom-up trace strength to be tokenised without requiring an enhancement from the blaster. However, if the blaster does fire, we observe increased temporal variance in the blaster activation dynamics, as reflected by attenuation in the grand average virtual N2pc and increased trial-to-trial variation in the virtual N2pc ERPimage. This, in turn, has a knock-on effect on T2's tokenisation process and its virtual P3, manifesting as attenuation in the grand average and increased variability in the virtual ERPimage. Quite a few targets, however, have insufficient strength to survive the delay in the blaster response. These targets are eventually missed, because of which the $ST^2$ model produces low simulated behavioural accuracy scores for targets during the AB.

## 7.4 Conclusion

In this chapter, we present human ERP evidence arguing in favour of increase temporal variance in the deployment of attention and subsequent working memory encoding during the AB. In Section 7.3.2, the ERPimages provide qualitative evidence for our hypothesis, which is supported quantitatively by the statistical analysis of the phase distributions. Our results suggest that the phase distribution for targets presented inside the AB is significantly less concentrated than the distribution of phases extracted for trials in which targets are presented outside the AB. The results from our analysis thus suggest that the attenuation of the grand average P3 and N2pc ERP components observed in previous studies is likely

to be due to increased temporal variance between the single trials of the raw EEG.

As the notion of increased temporal variation in the deployment of attention and working memory encoding is inherent to the theoretical framework of the $ST^2$ model, such effects are visible in the virtual ERPimages. In correlating virtual and human ERPs, we have shown that the $ST^2$ model replicates the human data in a qualitative manner. The EEG results thus provide support for the hypothesis of the $ST^2$ model.

# Chapter 8

# The attentional blink modulates the influence of target strength on conscious perception

This chapter investigates how target strength and the availability of attention affect target perception in RSVP. We show how behavioural accuracy scores for target letters belonging to the easy and hard categories (as defined in Chapter 6) differ significantly for both targets outside and inside the AB. When we extend this analysis to the P3 component, however, we find that 'easy-hardness' of targets affects the P3 for targets presented outside the AB but does not influence the P3 evoked by targets presented inside the AB. As the $ST^2$ model cannot account for these findings, we describe a modified theory that proposes two phases[1] of target perception in RSVP. Phase 1 determines whether the target can be behaviourally reported and is strongly sensitive to target strength. Phase 2 is only weakly sensitive to target strength, but influences the profile of the P3 component. We show how this *two-phase strength sensitivity theory* accounts for the experimental results presented in this chapter. Finally, the two-phase strength sensitivity theory predicts that T2s presented at lag 1 should show an 'easy-hard effect' both on behavioural accuracy and the P3 and this prediction is validated using experimental data.

[1] not to be confused with the two-stage theory of Chun and Potter (1995)

## 8.1 Introduction

Many of the cognitive operations we perform in our daily life, such as distributing attention or making decisions, have become established research areas in the field of cognitive neuroscience. The underlying cognitive mechanisms are studied using behavioural experiments and the neural correlates are identified by means of neuroimaging, electrophysiology and brain lesion studies.

The scientific study of consciousness, however, remains more controversial. How we become consciously aware of something is as much a psychological as it is a philosophical question. Why is it that only some types of behaviour require consciousness, whereas other brain activity remains completely unconscious? And even if we can define the requirements for a stimulus to be consciously perceived, it remains questionable whether one person perceives that stimulus in the same way as another. This is known as the *hard problem*, which describes the missing link between the objective world we live in and the subjective world we experience (Chalmers, 1996) and imposes the interesting theoretical question of 'whether two physically identical brains will have the same conscious state' (Koch, 2007). In any case, consciousness is often seen as a 'private experience' (Crick & Koch, 1995) and thus its scientific study is rather different from research in traditional areas of physics or biology (Koch & Hepp, 2006). For these reasons, the study of consciousness was, for a long time, not accepted as an area of the brain sciences, but delegated to philosophers and theologians. Since the advent of cognitive neuroscience, however, where researchers have started to bridge the gap between neural processes in the brain and cognitive phenomena traditionally studied by psychologists, the scientific study of consciousness has regained popularity.

### 8.1.1 The influence of bottom-up strength and attention on perception

A first requirement for an item to be consciously perceived is sufficient bottom-up strength. Under normal viewing conditions, most stimuli are strong enough to be consciously perceived. However, if stimulus representations are fleeting, as, for instance, in RSVP paradigms (see Section 2.3.1), they will sometimes be too weak to enter consciousness. Hence, the less active a neural representation of a stimulus is, the smaller the likelihood of that stimulus

entering awareness (Kanwisher, 2001).

However, this first point cannot be the sole requirement for conscious perception, as stimuli, which are equally strong in perceptual terms, in some cases succeed but in others fail to enter consciousness (Luck et al., 1996; Rees et al., 2000). Hence, the neural representation of a stimulus also needs to be accessible by other brain areas before it can enter awareness. The mechanism that creates this link is attention, which is likely to be a functional state of the brain (Baars, 1988). Koch and Tsuchiya (2007) have argued that attention and consciousness 'are distinct phenomena that need not occur together and can be manipulated using distinct paradigms.' To investigate the influence of bottom-up target strength and attention on conscious perception, we thus require an experimental paradigm where the availability of attention and bottom-up strength can be manipulated independently (Kim & Blake, 2005) and the AB is such a paradigm (see Section 2.3.2).

### 8.1.2 The P3 as a correlate of conscious perception

As discussed in Section 2.1.1, a P3 component is only evoked by those targets in RSVP that can be correctly reported (e.g. Kranczioch et al., 2003). Target items that are missed do not evoke a P3 component. The P3 is thus generally seen as an EEG correlate of encoding items into working memory (Vogel et al., 1998) and, by the same logic, a number of studies have proposed that the P3 serves as an index of conscious perception (e.g. Sergent et al., 2005; Kranczioch et al., 2007). In line with these previous studies, we use the P3 component as a correlate of conscious perception for targets in RSVP.

### 8.1.3 Overview

In the following chapter, we investigate the influence of bottom-up target strength and the availability of attention on the conscious perception of target items in RSVP. As discussed in section 8.1.1, the AB provides an experimental paradigm where - depending on the availability of attention - stimuli with equal bottom-up strength are sometimes seen and sometimes missed.

In the experiments that were conducted for this thesis, we did not externally vary stimulus strength (for instance by manipulating contrast). However, the analysis in Chapter 6

suggests a relationship between the identity of a target letter and behavioural accuracy as well as P3 size. The intrinsic stimulus characteristics (i.e. the shape of a particular target letter) thus allow us to classify target letters as belonging either to the easy or the hard target category, which in turn provides us with an *indirect* measure of target strength.

However, our notion of target strength relies completely on a significant difference in behavioural accuracy between easy and hard targets. Consequently, we have to show that, for each of the conditions of interest, there is a significant difference in behavioural accuracy scores between the easy and hard target group before we can perform any further analysis of the EEG data. Target letters are classified as belonging to the easy (respectively hard) category based on the data published in Bowman and Wyble (2007), using the methodology described in Chapter 6.

## 8.2 Methods

The analyses presented in this chapter are based on behavioural and EEG data from Experiments 1 and 2. The data for the *single target in RSVP* (target outside the AB) condition is taken from Experiment 1, whereas the *T2 following T1 at lag 3* (target inside the AB) and *T2 following T1 at lag 1* conditions are from Experiment 2. Please refer to the appendix for a detailed overview of the methods employed in these experiments. In the following, we describe the methods specific to the analyses presented in this chapter.

For the analyses of this chapter, the continuous data are segmented by extracting a time window of -200ms to 1000ms for the single target and the *T2 following T1 at lag 1* condition, whereas -500ms to 1000ms window is used for the *T2 following T1 at lag 3* condition[2]. The segmented EEG data are detrended to remove direct current drift artifacts. The single target and the *T2 following T1 at lag 1* data are baselined to the -200ms to 0ms window preceding target presentation and the data for *T2 following T1 at lag 3* are baselined to the -500ms to -300ms window with respect to target presentation (or the -200ms to 0ms period before the onset of the previous target, i.e. the T1).

After artifact rejection, each of the conditions contain the following number of trials.

[2]The segmentation window for the *T2 following T1 at lag 3* starts at -500ms to allow us to baseline to the 200ms period before the onset of the preceding target (i.e. the T1).

For the *single target in RSVP* (target outside the AB): Easy-Correct contains 1069 trials, Hard-Correct contains 939 trials, Easy-Incorrect contains 100 trials and Hard-Incorrect contains 225 trials. For the *T2 following T1 at lag 3* (target inside the AB): Easy-Correct contains 511 trials, Hard-Correct contains 358 trials, Easy-Incorrect contains 269 trials and Hard-Incorrect contains 436 trials. For *T2 following T1 at lag 1*: Easy-Correct contains 521 trials, Hard-Correct contains 430 trials, Easy-Incorrect contains 44 trials and Hard-Incorrect contains 118 trials.

Before plotting, the trials within each condition are low-pass filtered at 25Hz, shuffled to intermix the data of all subjects and then vertically smoothed using a sliding window of 40 trials to increase the visual signal-to-noise ratio. The statistical analyses are performed on the subject averages of each condition using the methodology described in the appendix.

## 8.3 Results

### 8.3.1 Behavioural

The prerequisite for the following EEG analysis is a significant difference in behavioural accuracy scores between easy and hard letters for both targets presented outside and also inside the AB. Using the method from the easy/hard analysis of Chapter 6, target letters are classified as being easy or hard based on a previously published study (Bowman & Wyble, 2007). Accordingly, when analysing the behavioural data from Experiment 1, target letters T, K, U, V, L, D and G are categorised as easy, whereas E, C, B, P, F, J and R belong to the hard category. For the analysis of Experiment 2, target letters T, K, U, V, L, D, G, N and H are categorised as easy, whereas E, C, B, P, F, J, R, Y and A belong to the hard category. The difference in the number of letters per category is because Experiment 2 contained four additional target letters (see appendix for details), hence, both the easy and hard condition contain nine letters (and not seven letters as in Experiment 1). However, to re-emphasise, the subdivisions in Experiment 1 and also Experiment 2 are inherited from Bowman and Wyble (2007) and are thus a priori.

If we apply this analysis to the *T2 following T1 at lag 3* condition, the accuracy scores for targets belonging to the easy and hard categories (as defined in the previous paragraph) are 66% (SEM 4) and 46% (SEM 5), respectively. The difference is highly significant;

$F(1,17) = 59.4$, MSE $< 0.01$, $p < 0.001$. The *T2 following T1 at lag 3* condition from Experiment 2 is thus used to investigate target processing inside the AB.

In order to investigate the processing of targets outside the AB, we can employ the *T1 with T2 presented at lag 8* condition from Experiment 2 or the single target in RSVP from Experiment 1. The easy/hard analysis for the *T1 with T2 presented at lag 8* condition, however, does not meet the requirements for further EEG analysis. The difference between easy and hard targets in the *T1 with T2 presented at lag 8* condition from Experiment 2 is only marginally significant; easy 87% (SEM 2) vs. hard 82% (SEM 3), $F(1,17) = 4.4$, MSE $< 0.01$, $p = 0.051$. This is likely to be due to ceiling effects, as T1 lag 8 accuracy is relatively high for both easy and hard target letters.

The *single target in RSVP* condition overcomes this problem, as we increased presentation rate to 50ms per item in Experiment 1 for target detection accuracy to be below ceiling. Single targets in RSVP show a highly significant effect of target difficulty, mean accuracy is 82% (SEM 4) for easy and 62% (SEM 4) for hard letters; $F(1,19) = 94.1$, MSE $< 0.01$, $p < 0.001$. Hence, we employ the single target condition from Experiment 1 to investigate target processing outside the AB.

### 8.3.2 EEG

We visualise the relationship between the human P3 component and target difficulty as well as accuracy by plotting ERPimages for targets outside and inside the AB. As discussed in Chapter 7, ERPimages allow only a qualitative inspection of the data. In order to perform a statistical analysis of our results, we extract the mean P3 size per subject for each accuracy-target difficulty combination for both targets outside and inside the AB.

#### Targets outside the AB

Figure 46A depicts the ERPimage for the P3 component evoked by targets outside the AB and trials are sorted by target difficulty and accuracy with respect to the target. The bar chart in Figure 46B shows that mean P3 size is influenced by our indirect measure of target strength for correctly reported targets presented outside the AB. Targets in the Easy-Correct condition (8.9$\mu$V, SEM 0.9) have a significantly larger P3 than targets in the Hard-Correct condition (6.7$\mu$V, SEM 1.0). For incorrectly reported targets the ERP

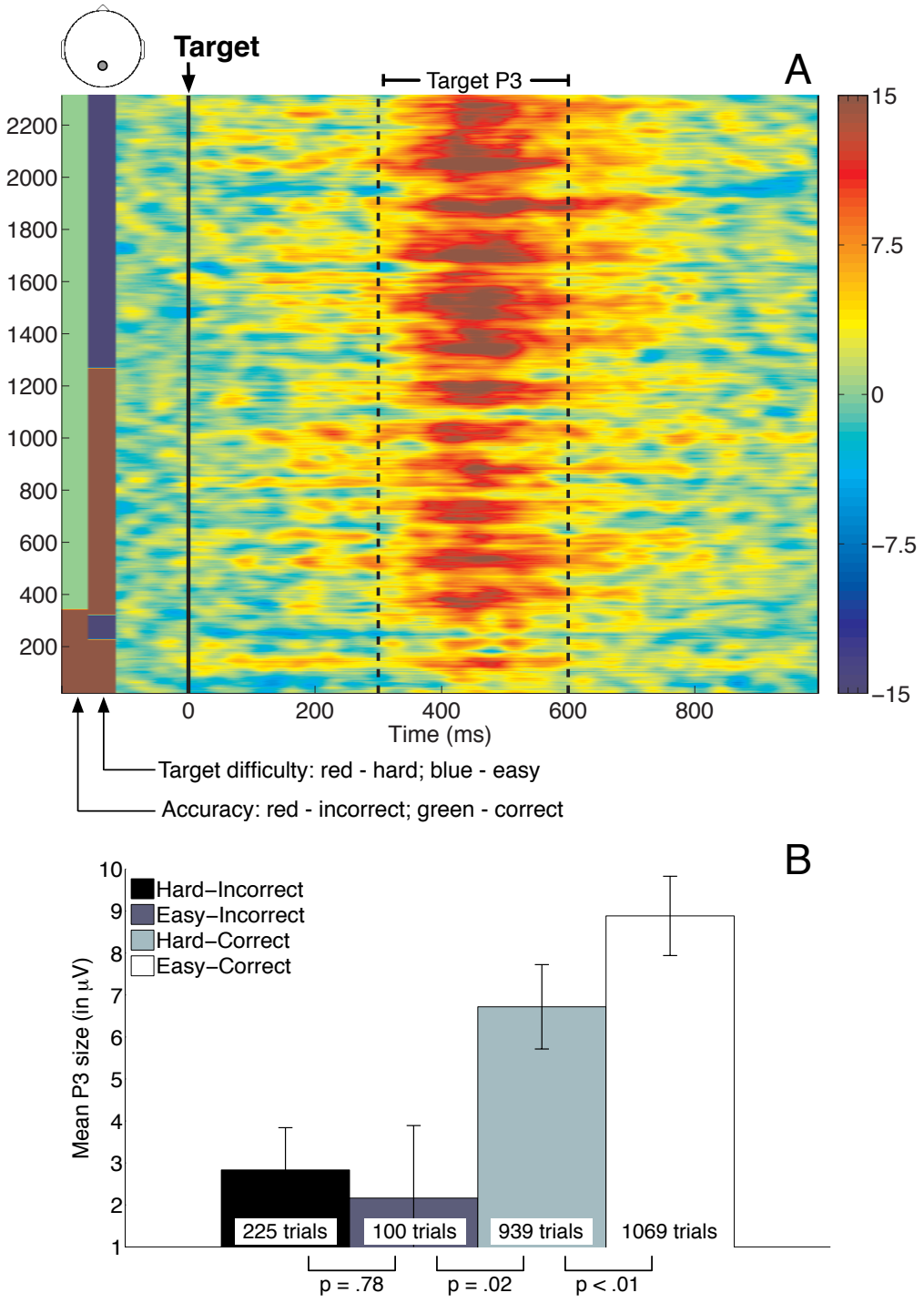


*Figure 46* Panel A: ERPimage from electrode Pz for a target outside the AB. Colour strips to the left of the plot indicate the accuracy and target difficulty for that particular trial. Trials within each accuracy/target difficulty category are randomly shuffled before vertically smoothing to average over individual subject differences. Panel B: Bar chart displaying the mean P3 size (300-600ms with respect to target onset) for each accuracy-target difficulty combination. The error bars depict the standard error of the mean.

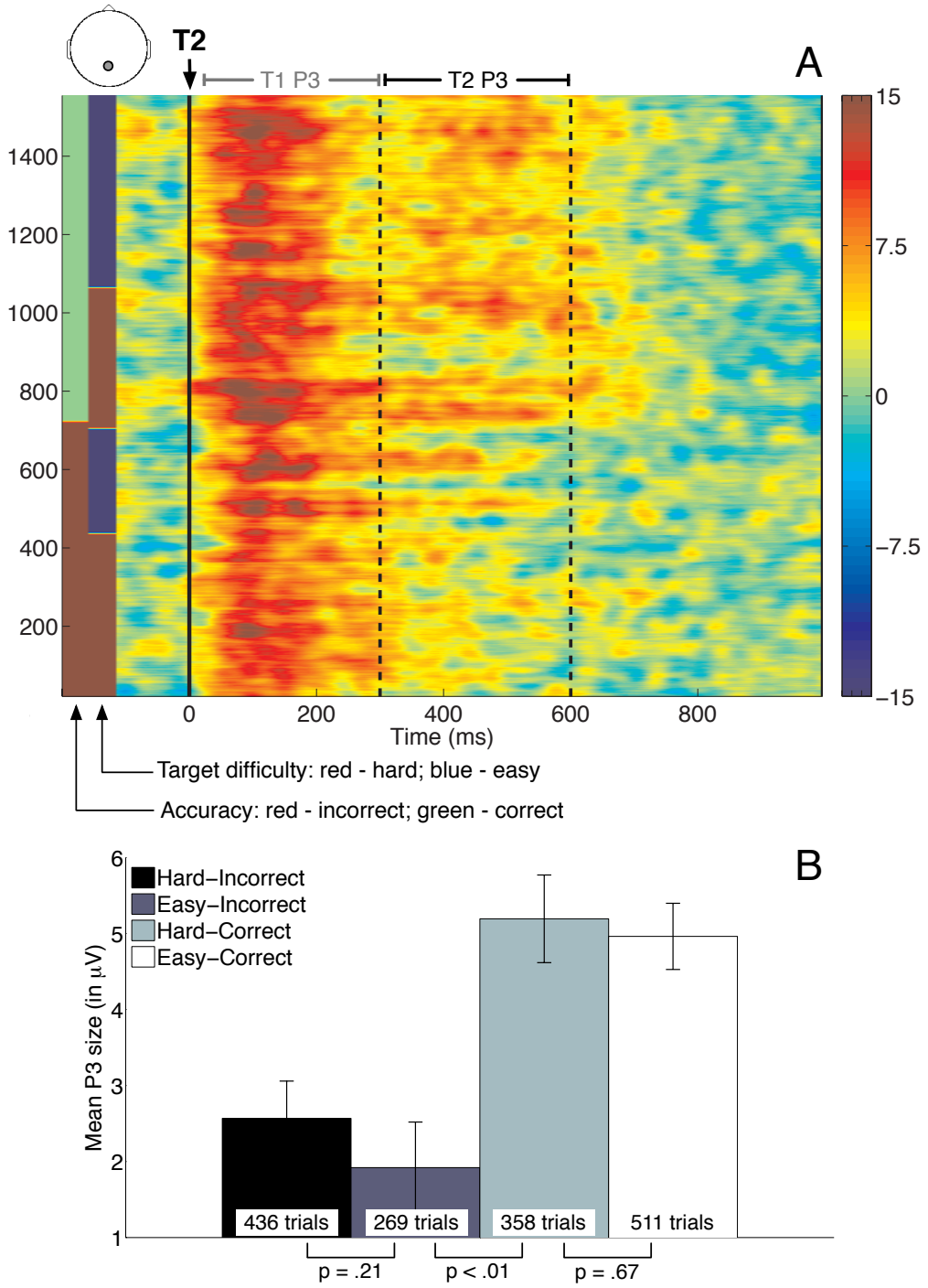


*Figure 47* Panel A: ERPimage from electrode Pz for targets presented inside the AB (T2 at lag 3 following a correctly identified T1). Colour strips to the left of the plot indicate the accuracy and target difficulty for that particular trial. Trials within each accuracy/target difficulty category are randomly shuffled before vertically smoothing to average over individual subject differences. Panel B: Bar chart displaying the mean P3 size (300-600ms with respect to T2 onset) for each accuracy-target difficulty combination. The error bars depict the standard error of the mean.

activity in the P3 window remains close to baseline; Hard-Incorrect (2.8$\mu$V, SEM 1.0) vs. Easy-Incorrect (2.2$\mu$V, SEM 1.7).

Individual pairwise comparisons show that target difficulty influences P3 size for correctly reported targets and the effect is significant; $F(1,19) = 28.2$, MSE = 1.5, $p < 0.001$[3]. There is a significant difference in P3 size between the Easy-Incorrect and Hard Correct conditions: $F(1,19) = 7.2$, MSE = 25.9, $p = 0.016$. However, for the incorrectly reported targets, strength has no effect and the difference in P3 sizes between the Easy-Incorrect and the Hard-Correct condition is not significant ($F(1,19) = 0.1$, MSE = 49.8, $p = 0.779$).

**Targets inside the AB**

Figure 47A depicts the ERPimage for the P3 component for targets inside the AB (*T2 following T1 at lag 3*) and is sorted by target difficulty and accuracy with respect to T2. The bar chart from Figure 47B depicts the mean P3 size (300-600ms with respect to T2 presentation) for each target difficulty and accuracy combination.

Individual pairwise comparisons show that - in contrast to targets outside the AB - target difficulty does not have a significant effect on P3 size for targets inside the AB. Both for correctly reported targets (Easy-Correct 5.0$\mu$V (SEM 0.4) vs. Hard-Correct 5.2$\mu$V (SEM 0.6); $F(1,17) = 0.2$, MSE = 2.5, $p = 0.664$) and incorrectly reported targets (Easy-Incorrect 1.9$\mu$V (SEM 0.6) vs. Hard-Incorrect 2.6$\mu$V (SEM 0.5); $F(1,17) = 1.7$, MSE = 2.2, $p = 0.209$) the difference between P3 sizes is not significant. The difference between the Easy-Incorrect and Hard-Correct conditions, however, is significant ($F(1,17) = 26.4$, MSE = 3.7, $p < 0.001$), which suggests that the high p-values in the previous analyses are not due to a lack of statistical power. Instead, this indicates that there is indeed no difference in P3 size between easy and hard targets both when they are correctly and incorrectly reported.

[3]There is slight inconsistency between the results of the Easy-Correct and Hard-Correct comparison reported here and the results presented in Chapter 6. This is due to differences in the detrending and filtering procedures. Note though that although the statistical analyses show slightly different values, in qualitative terms, the results are the same.

## 8.4 Discussion

This chapter presents EEG data investigating how bottom-up target strength and the availability of attention modulate conscious perception of targets in RSVP. As the AB provides a paradigm where target strength and the availability of attention can be manipulated independently, we compare the EEG signatures (and, in particular, the P3 component) of targets presented outside the AB to targets that are presented during the AB. In the first part of the discussion, we interpret our EEG results. Following this, we present virtual ERPs to illustrate the perspective of the $ST^2$ model on this issue. As we will show, however, the $ST^2$ model cannot account for the EEG results presented in this chapter. Hence, we speculate about a new theory that proposes an explanation for both our EEG results and also previous findings of an all-or-none bifurcation of behavioural visibility ratings (Sergent & Dehaene, 2004) and the P3 component (Sergent et al., 2005) during the AB (see also Section 2.3.2). Finally, we use this new theory to make an experimental prediction, which we test using our behavioural and EEG data.

### 8.4.1 Target difficulty affects the P3 for targets outside but not inside the AB

As shown in Section 8.3.1, there is a significant difference in accuracy scores between easy and hard targets both outside and inside the AB. Hence, in terms of behavioural accuracy, the identity of a target letter (i.e. whether it belongs to the easy or hard category of letters) has an influence on target detection both if the target is presented individually (i.e. outside the AB) and also if it is presented during the AB.

In Chapter 6, we showed that target difficulty affects the size of the P3 component for individually presented targets (i.e. targets presented outside the AB) that are correctly reported. We proposed that an easy target letter has more bottom-up strength than a hard letter and this increases the size of the P3 component evoked by the easy target. The results presented in Section 8.3.2 re-emphasise how, for targets outside the AB, 'easy-hardness' of targets affects the size of the P3 component.

In Section 8.3.2, we perform the same analysis for targets presented during the AB. From the behavioural analysis (Section 8.3.1), we know that intrinsic stimulus characteristics

(i.e. whether the target is easy or hard) affect target perception if the target is presented during the AB. Interestingly, however, 'easy-hardness' of targets does not influence the P3 component for targets during the AB. The bar chart in Figure 47B illustrates that there is no statistically significant difference in P3 size for easy and hard targets, both if the target is correctly and incorrectly reported. It seems that the P3 component is influenced by different factors depending on whether a target is presented outside or inside the AB, which is intriguing.

### 8.4.2 Virtual ERPs from the ST$^2$ model

In the previous chapters of this thesis, we have shown how we can visualise the theory underlying the ST$^2$ model by plotting virtual ERPs for the conditions of interest. Accordingly, we express the ST$^2$ model's theoretical standing on the effect of target strength on target perception outside and inside the AB by generating the corresponding virtual ERPs. In the human data, we have no direct measure of target strength and, hence, use target difficulty (i.e. whether a target is easy or hard) as an indirect measure of target strength. In the model, however, target strength is precisely defined by each target's strength value. Consequently, we generate virtual ERPimages (see Section 4.3.5 for a description of virtual ERPimage methodology) that are sorted by the target strength value and simulated accuracy on each trial.

#### Targets presented outside the AB

If a target is presented to the ST$^2$ model individually (i.e. outside the AB), the blaster is available. Consequently, the blaster enhances the item's type representation as soon as this item has been identified as a target by the task filter. Whether a target is successfully tokenised depends on its bottom-up strength. After receiving an enhancement from the blaster, targets outside the AB will normally have enough bottom-up strength for their type representation to bind to a token and thus they will be encoded into working memory. Very weak targets, however, fail to gather sufficient strength and cannot initiate tokenisation. Such items fail to bind to a working memory token and are not 'seen' by the ST$^2$ model.

Figure 48 shows the activity underlying the virtual P3 for a single target in RSVP. The colour strips on the left indicate the target's accuracy and strength, respectively, for the

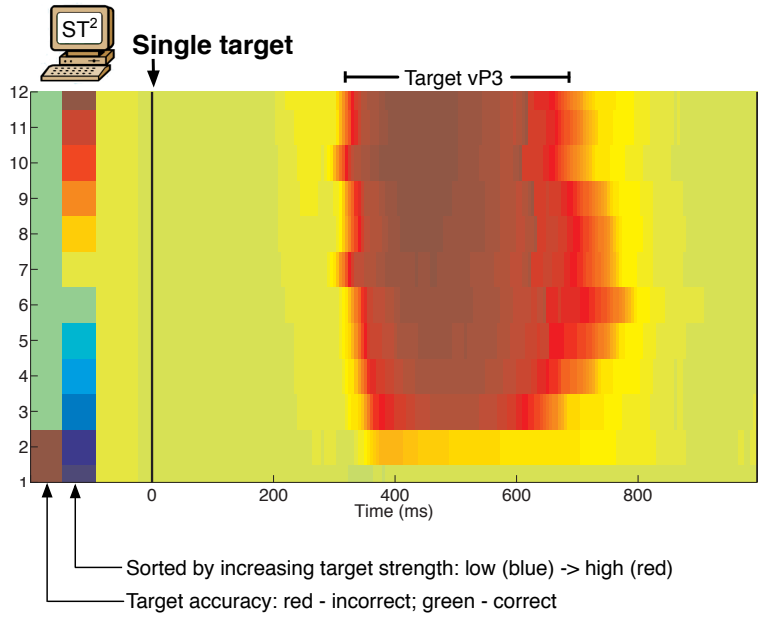


*Figure 48* Virtual P3 from the $ST^2$ model for targets outside the AB (a single target in RSVP). Colour strips to the left of the plots indicate the accuracy and target strength for that particular trial of the simulation.

corresponding trial of the simulation. The virtual ERPimage is sorted by target strength, from the lowest strength value at the bottom to the highest at the top of the plot. The virtual ERPimage shows how in the $ST^2$ model, targets have to overcome a critical strength value before they are able to initiate tokenisation. Whereas targets with strength values below the threshold are 'missed', targets above the threshold are 'seen' by the $ST^2$ model.

### Targets presented inside the AB

As discussed in Chapter 7, the duration of T1's tokenisation is determined by T1's strength. The blaster is suppressed while T1 is tokenised, hence, the availability of attention for T2 during the AB depends on how long it takes to tokenise T1. For targets presented inside the AB, successful tokenisation thus depends not only on T2's strength, but also on the strength of the preceding T1 and consequently the availability of attention.

Figure 49 illustrates this issue in the virtual P3 ERPimage for a T2 presented at lag 3 following a correctly identified T1, i.e. a target inside the AB. Again, this plot displays T2 accuracy and strength to the left of the figure and trials are sorted by T2 target strength.

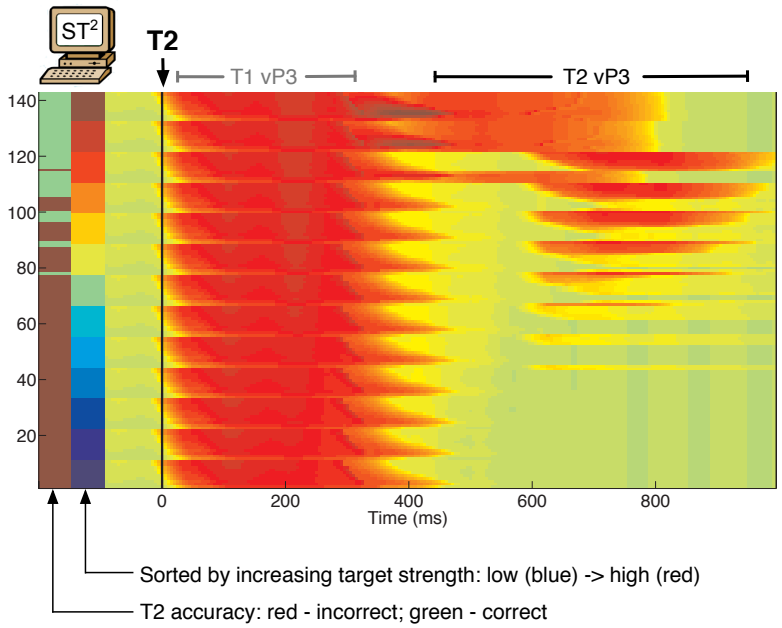


*Figure 49* Virtual P3 from the $ST^2$ model for targets inside the AB, i.e. a T2 presented at lag 3 following a correctly reported T1. Colour strips to the left of the plots indicate the accuracy and target strength for that particular trial.

Furthermore, within each T2 strength value, the trials are sorted in ascending order by T1's target strength values. In contrast to targets outside the AB, the blaster is suppressed by T1's tokenisation process during the AB. As target representations decay over time, many targets with lower strength values (trials 1-80 in Figure 49) fail to 'outlive' the unavailability of attention and cannot be tokenised. Even targets with medium strength values (trials 60-100 in Figure 49) are mostly 'missed' during the AB. These trials show some marginal virtual P3 activity, which is due to activity in higher layers of stage one contributing towards the virtual P3. Only a few medium strength targets, where T1 is tokenised particularly quickly and the blaster becomes available earlier, are able to initiate the tokenisation process and are 'seen' by the $ST^2$ model. Targets with high strength values (displayed at the top end of the ERPimage in Figure 49) have sufficient bottom-up strength to be tokenised despite being presented during the AB.

**The ST$^2$ model fails to explain the results for targets during the AB**

For targets outside the AB, the higher the target strength value, the greater the likelihood that the target will be correctly reported by the ST$^2$ model. Hence, target strength directly affects simulated behavioural accuracy. Furthermore, virtual P3 size increases with larger target strength values. A target strength value is the model's equivalent of a target's intrinsic stimulus characteristics in the human data (i.e. whether it is easy or hard; see also Chapter 6). Consequently, in the ST$^2$ model, whether a target is easy or hard influences both simulated accuracy and the virtual P3. For targets outside the AB, the model is thus in line with the human results.

In most cases, target detection during the AB depends on the availability of attention, which, in turn, is determined by the amount of time it takes to process the preceding target (i.e. the T1). Consequently, there are some trials where a stronger T2 is 'missed' because T1 processing takes too long and, vice versa, other trials where weaker T2s are 'seen' because T1 is tokenised quickly. Nevertheless, target detection and especially the virtual P3 for targets during the AB (see Figure 49) are also strongly influenced by target strength, i.e. targets with higher strength values generate larger virtual P3 components. This is in contrast with the human data from Section 8.3.2, where, during the AB, the P3 component seems unaffected by target strength.

### 8.4.3 The two-phase strength sensitivity theory

The ST$^2$ model cannot adequately explain the experimental results presented in Section 8.3. Consequently, and in order to interpret the theoretical contribution of the results, we propose a modified theory that attempts to account for this chapter's findings. To reiterate, the data that we are trying to explain suggests that 'easy-hardness' of targets affects behavioural accuracy scores both for targets inside and outside the AB. The P3 component, however, is only influenced for targets outside the AB.

Our proposed *two-phase strength sensitivity theory*[4], is based on theoretical concepts borrowed from the ST$^2$ model, however, deviates from the ST$^2$ model in a number of ways. Note that rather than being a complete formal model, the two-phase theory is currently

[4]referred to as *two-phase theory* from here

only an informal hypothesis.

### Description of the two-phase strength sensitivity theory

The two-phase theory describes the process of target perception in RSVP. We do not describe early stages of visual processing, but focus on later stages of processing where targets are encoded into working memory (i.e. stages of target processing that would involve the task filtered layer in the $ST^2$ model). Neural activation from these stages of target processing is assumed to contribute to the P3 component.

Figure 50 depicts the activation traces for varying target strengths according to the two-phase theory. After target presentation, the target's activation trace has to overcome a threshold in order for it to be consciously perceived. In the following, we describe the two phases of target perception, as proposed by the two-phase theory.

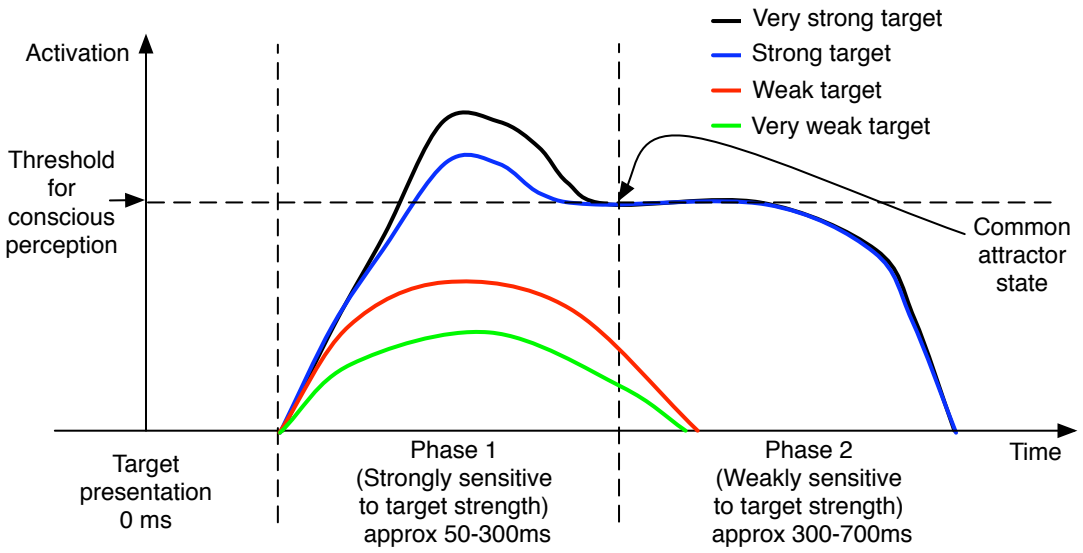


*Figure 50* The two-phase theory's hypothesised target activation traces for varying target strengths. The figure describes the profile of activation without enhancement from the blaster.

**Phase 1 - strongly sensitive to target strength** Phase 1 is strongly sensitive to target strength. Hence, in phase 1, targets with different strength values have different activation profiles. The activation of a target during phase 1 is the critical factor influencing behavioural accuracy scores, as a target's activation level has to overcome a critical threshold

(marked by the dotted line in Figure 50) to become eligible for conscious perception. Weak targets without sufficient bottom-up strength decay before their activation level reaches the threshold, which means that these targets cannot be behaviourally reported.

**Phase 2 - weakly sensitive to target strength** Phase 2 succeeds phase 1 and is only weakly sensitive to target strength, in the sense that there are two possible activation levels: a) a steady-state activation level or b) zero activation. Targets with enough activation during phase 1 will have passed the threshold when phase 2 occurs, which means that their activation remains in a steady-state.

**Attentional enhancement from the blaster** The blaster provides attentional enhancement to targets. Similarly to the concept described in the $ST^2$ model, the blaster fires once an item has been identified as a target by the system. The enhancement from the blaster increases targets' activation levels to facilitate successful working memory encoding.

As initial support for this hypothesis, an ERP study by Del Cul et al. (2007) has indeed identified two phases of target processing, which have different sensitivities to target strength. Although not an RSVP study, Del Cul et al. (2007) manipulate target strength using masking. Importantly, they find an early phase that is highly sensitive to masking strength and a later phase, which - although not as weakly sensitive as we are proposing - is certainly a lot less sensitive than phase 1. This is most evident in Figure 8 of Del Cul et al. (2007), in particular, the panel depicting ERP activity localised to posterior ventral temporal sources of the brain.

### Targets outside the AB

**Activation traces** The hypothesised activation traces for targets presented outside the AB are depicted in Figure 51. As previously discussed, a target outside the AB is presented individually. Consequently, the blaster is available to enhance the target's representation as soon as it is detected by the system. For targets outside the AB, the blaster enhancement occurs during phase 1, which is strongly sensitive to target strength. As seen in Figure 51, the blaster proportionally increases target activation, which then remains at a steady-state activation level. The activation level of this steady-state differs between targets, as it is

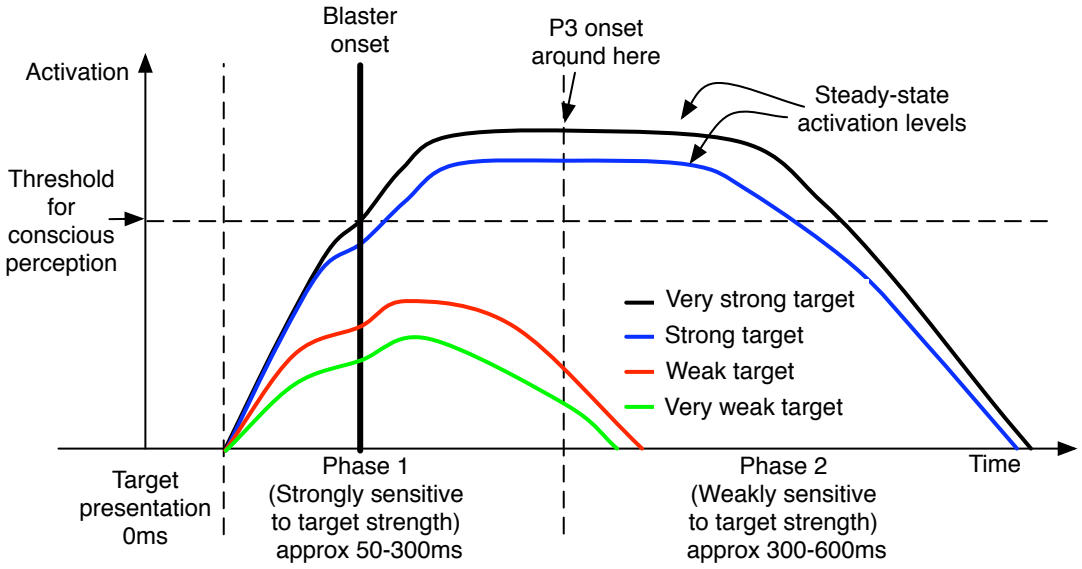


*Figure 51* Target activation traces from the two-phase theory including blaster enhancement. When targets are presented individually (or outside the AB), the blaster is available and fires as soon as the target has been detected by the system, i.e. during phase 1. For strong targets, the enhancement from the blaster leads to a steady-state activation level where the amount of activation corresponds to the target's initial strength value. Weak targets do not have sufficient activation strength to fire the blaster and their activation decays back to baseline during phase 1.

determined by the target's initial strength value. Weak targets, however, fail to fire the blaster. Consequently, the activation for these targets decays to baseline during phase 1.

**Easy-hard affects behavioural accuracy and the P3** Phase 1 is critical for determining whether a target can be behaviourally reported, as it is in phase 1 that target activation has to overcome the critical threshold to remain active throughout phase 2. For targets outside the AB, blaster enhancement occurs during phase 1 and the activation trace of a target is increased proportionally to its previous strength (see Figure 51). Consequently, the higher a target's strength, the higher the probability of that target entering a steady-state activation level above the critical threshold, which allows the target to be behaviourally reported. This direct relationship between target strength and the likelihood of target detection accounts for the easy-hard effect in our behavioural accuracy for targets outside the AB (see Section 8.3.1).

The P3 component is hypothesised to reflect the activation level of a target during working memory encoding. We hypothesise the P3 component to onset around the start of

phase 2. As seen in Figure 51, the two-phase theory proposes that targets' activation levels differ depending on their initial strength value. For targets outside the AB, the size of the P3 component is determined by a target's activation level during the period of steady-state activation, which is in turn determined by its initial bottom-up strength. This accounts for the easy-hard effect on P3 component size in our EEG results from Section 8.3.2.

Furthermore, the two-phase theory can account for the results from the previously mentioned ERP study that investigates the effect of masking on the P3 component. Del Cul et al. (2007) find that the size of the P3 evoked by individually presented targets varies depending on the SOA between the target and the following mask (main effect: $p < 0.001$). It can be assumed that the shorter the SOA between the mask and the target, the more strongly the target's representation is weakened through masking. Shorter SOAs can be associated with lower target strength and, vice versa, the longer the SOA between target and mask, the higher the target strength. Del Cul et al. (2007) find that P3 size increases linearly with increasing target strength ($p < 0.001$), hence, this is further evidence for target strength affecting P3 size when targets are presented individually (which, if we extrapolate to our experimental setup, would correspond to targets presented outside the AB).

### Targets inside the AB

**Activation traces** The hypothesised activation traces for a target presented inside the AB are depicted in Figure 52. In line with the $ST^2$ model, the two-phase theory suggests that the blaster is suppressed while T1 is encoded into working memory. Hence, during the AB, T2 is presented before T1's working memory encoding process has completed. Consequently, the onset of the blaster is delayed for a target presented inside the AB[5] and, as illustrated in Figure 52, does not occur until phase 2, which is only weakly sensitive to target strength.

Without blaster amplification, weak targets decay to baseline during phase 1 and, consequently, do not have enough strength to fire the blaster once it becomes available during phase 2. Hence, weak targets presented during the AB show no activation during phase 2. Strong targets presented during the AB, on the other hand, have enough strength to

[5]Indeed, stepping beyond the $ST^2$ model, there is a good deal of evidence that T2 consolidation is delayed during the AB (Vogel & Luck, 2002; Sessa et al., 2006).

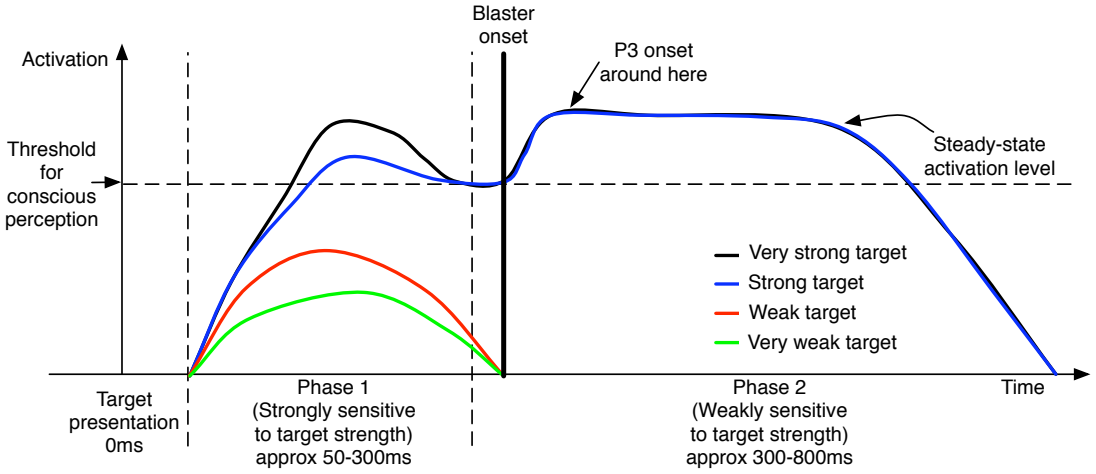


*Figure 52* The two-phase theory's hypothesised target activation traces including blaster enhancement. For a target (T2) that is presented during the AB, the blaster is delayed and does not fire until phase 2, as the blaster is suppressed while the preceding target (T1) is encoded into working memory. Due to the delay, only strong targets can fire the blaster, weak targets decay to baseline during phase 1. Note that the figure only contains the activation trace for the T2; T1 activation is not shown.

overcome the threshold for conscious perception and enter a common attractor state by the end of phase 1. These targets are in attractor states with equal activation levels when the blaster gets fired, which proportionally increases the target's activation. Hence, the activation for all targets that do manage to fire the blaster during the AB is elevated to an equal steady-state activation level during phase 2.

To summarise, weak targets do not benefit from the blaster and show no activation, whereas strong targets receive the blaster enhancement whilst in an attractor state with equal activation levels and, hence, show equal levels of steady-state activation irrespective of their initial bottom-up strength value. The two-phase theory thus proposes that for targets presented during the AB, activation levels in phase 2 are all-or-none.

**Easy-hard affects behavioural accuracy but not the P3** The probability of a target being behaviourally reported is determined in phase 1. For targets inside the AB, however, the blaster does not fire until after phase 1. Hence, behavioural accuracy is not influenced by the blaster but, instead, is determined by a target's initial strength. The stronger a target, the greater the likelihood of that target's activation overcoming the threshold for

conscious perception. Consequently, for targets inside the AB, there is an easy-hard effect on behavioural accuracy scores and this is in agreement with the results from Section 8.3.1.

By the time the blaster fires for targets inside the AB, targets with sufficient strength have already entered a common attractor state, whereas weak targets have failed to reach the threshold and have decayed to baseline (as depicted in Figure 52). Our EEG results from Section 8.3.2 suggest that for targets inside the AB, the P3 component is not influenced by target strength. We can thus use the two-phase theory to propose two alternative explanations for the lack of an 'easy-hard effect' on the P3 during the AB.

According to the first hypothesis, it might be that only activation occurring after blaster onset contributes to the P3 component. For targets during the AB, the blaster does not fire until phase 2, so phase 1 does not contribute towards the P3. Only phase 2 contributes to the P3 component and the activation in phase 2 is all-or-none, in that the activation for strong targets converges at a steady-state level whereas the activation for weak targets decays back to baseline. However, the P3 is an ERP component that can be recorded from various scalp locations (see Section 2.1.1 for details). Hence, for phase 1 activation not to contribute towards the P3 component, phase 1 activation would have to originate from brain regions that are spatially (and maybe also temporally) separate from phase 2 activation (presumably restricted to occipital regions) or indeed to be very weak compared to later activation from phase 2.

The second hypothesis states that phase 1 activation might be partially visible in the P3 component. However, for a target presented during the AB (i.e. T2 at lag 3), phase 1's contribution to the P3 is overlaid by the P3 component evoked by the preceding target (i.e. T1). Consequently, phase 1 activation related to the T2 would not be visible in the P3 component evoked by that T2. Only phase 2 activation (which shows an all-or-none profile for targets presented during the AB) would be visible in the P3 component and this would account for the results from Section 8.3.2 showing that the P3 evoked by targets presented inside the AB is unaffected by target strength. As it stands, we cannot distinguish between these two alternative hypotheses, as they could both account for the data.

**The two-phase theory predicts conscious perception to be all-or-none during the AB**

As discussed in Section 2.3.2, there is behavioural evidence showing that participants' visibility is bimodal during the AB (Sergent & Dehaene, 2004). Observers were asked to report the extent to which the target had been perceived using a visibility scale ranging from 'Nothing' (0%) to 'Maximum visibility' (100%). For targets inside the AB, the majority of responses were concentrated around the minimum and maximum of the visibility scale (see Figure 53A). For targets presented outside the AB, however, the responses were gradually distributed with no clear threshold in visibility rankings (see Figure 53B).

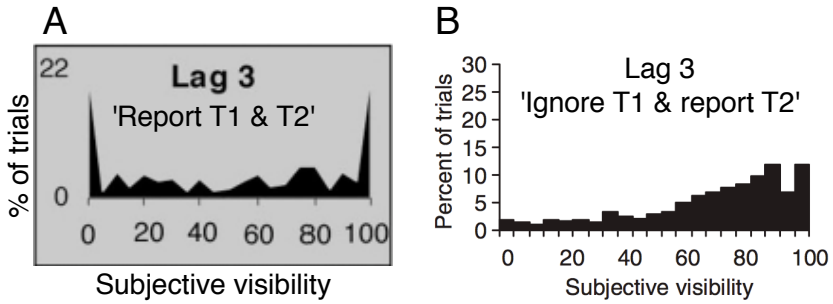


*Figure 53* Panel A: Response distribution (percentage of trials in each visibility category) for a T2 presented during the AB, i.e. T2 follows T1 at lag 3 and participants are instructed to report both T1 and T2. Adapted from Sergent and Dehaene (2004). Panel B: Response distribution (percentage of trials in each visibility category) for a T2 presented outside the AB, i.e. T2 follows T1 at lag 3 but participants are instructed to ignore T1. Adapted from Sergent et al. (2005).

With respect to the visibility ratings, the P3 component was also found to be distributed in an all-or-none fashion during the AB (Sergent et al., 2005). Trials with higher visibility scores showed a large P3 component, whereas trials with low visibility scores showed virtually no P3 component (see Figure 54).

Sergent and colleagues' results cannot be directly related to our data, as Sergent et al. used their visibility ranking, whereas our experiment employs the indirect measure of target difficulty to index target strength. However, like Sergent et al. (2005), we find that the P3 component is unaffected by target strength (i.e. showing an all-or-none pattern) when targets are presented inside the AB. For targets outside the AB, we find that the P3 varies with target strength, which is in agreement with the gradually distributed behavioural

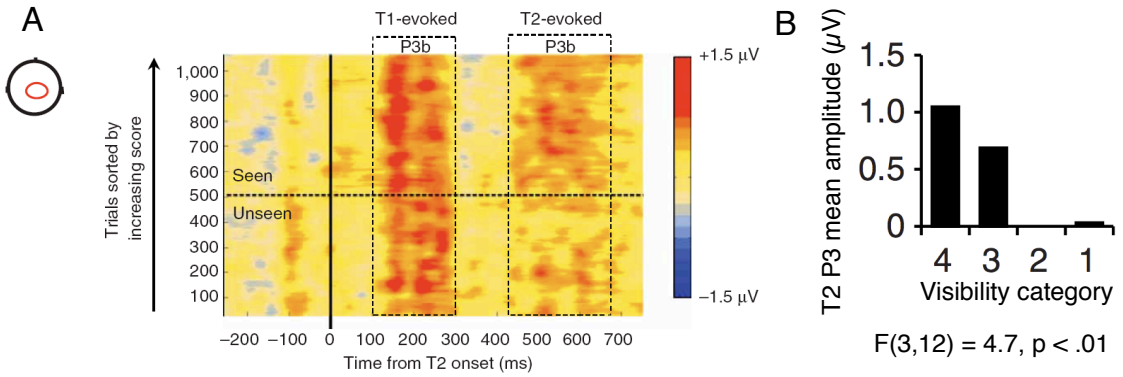


*Figure 54* Panel A: ERPimage plot of the P3 components evoked by T1 and T2 when T2 is presented during the AB (at lag 3). Trials are sorted by visibility score, lowest visibility at the bottom. Voltage values (positive in red and negative in blue) are smoothed over 50 trial windows. Panel B: Histogram of mean amplitude of the T2 P3 component (presented during the AB) per visibility category. Categories 4 and 3 correspond to visibility scores $> 50\%$ ('seen' trials), whereas categories 2 and 1 to visibility scores $< 50\%$ ('unseen' trials). Within seen and unseen trials, the categories are classified per participant using the median of that participant's response distribution. Adapted from Sergent et al. (2005).

responses from Sergent et al. (2005). If we assume that visibility rating is governed by phase 2 activation strength (which would be the natural interpretation), then the two-phase theory explains Sergent et al.'s findings.

Sergent and colleagues argue that conscious perception is all-or-none when perception is determined by the availability of attention, as is the case during the AB. We argue though that it is not the absolute unavailability of attention that causes the all-or-none pattern. Rather, it is the delay of the blaster firing that causes all-or-none perception during the AB (and consistent with this hypothesis, we know that T2 consolidation is indeed delayed during the AB (Vogel & Luck, 2002; Sessa et al., 2006)). Specifically, the two-phase theory proposes that, due to the delayed firing of the blaster during the AB, the blaster's enhancement of target activation does not have its effect until phase 2, which is only weakly sensitive to target strength. In phase 2, targets have either entered the common attractor state required for perception or their activation has decayed back to baseline, hence, activation traces show an all-or-none profile.

For targets outside the AB, however, the two-phase theory suggests that the blaster enhancement occurs during phase 1, which is sensitive to target strength. Hence, the

strength of the percept (which is assumed to correspond to the level of the steady-state activation plateau) varies and one would expect a graded continuum of conscious perception as reported in Sergent et al. (2005).

**Prediction: Target difficulty affects behaviour and P3 for a T2 at lag 1**

The two-phase theory, as described in the previous sections, suggests the following prediction for a T2 following T1 at lag 1. Because at lag 1, T2 is presented in immediate succession to T1, T2 will receive a major part of the blaster enhancement, which was initially intended for T1 (see also Chapter 6). This enhancement occurs very early during phase 1 of T2 processing, which - as previously described - is strongly sensitive to target strength. Consequently, the two-phase theory predicts that, in a similar fashion to targets outside the AB, T2s presented at lag 1 should show an 'easy-hard effect' both in terms of behavioural accuracy scores and also P3 size.

**Behavioural results** T2s that are presented at lag 1 and belong to the easy target category have a mean accuracy score of 90% (SEM 2), whereas hard T2s at lag 1 have an average accuracy of 78% (SEM 3). The difference is significant: $F(1,17) = 12.3$, MSE = 0.01, $p = 0.003$. There is an 'easy-hard effect' on behavioural accuracy and, hence, the experimental data are consistent with the prediction from the two-phase theory.

**EEG results** Figure 55 depicts the ERPimage for the P3 component for a T2 presented at lag 1. Trials are sorted by target difficulty and accuracy with respect to T2. Since this is a lag 1 case, there is one joint P3 component for both T1 and T2 (see Chapter 6 for an extensive discussion of this issue). In a similar fashion to the analysis of targets outside the AB, the T2 at lag 1 analysis suffers from a lack of trials in conditions where the target is incorrectly reported. For correctly reported targets, however, there are sufficient number of trials to extract the mean P3 size for each accuracy and target difficulty combination.

Figure 56 shows the mean P3 size for each of the target difficulty and accuracy combinations in the T2 at lag 1 condition. As discussed in Chapter 6, there is one joint P3 component for T1 and T2 at lag 1. Consequently, the separation of P3 activation associated with each of the targets (T1 and T2) is difficult and T2 'easy-hard effects' on the P3 might

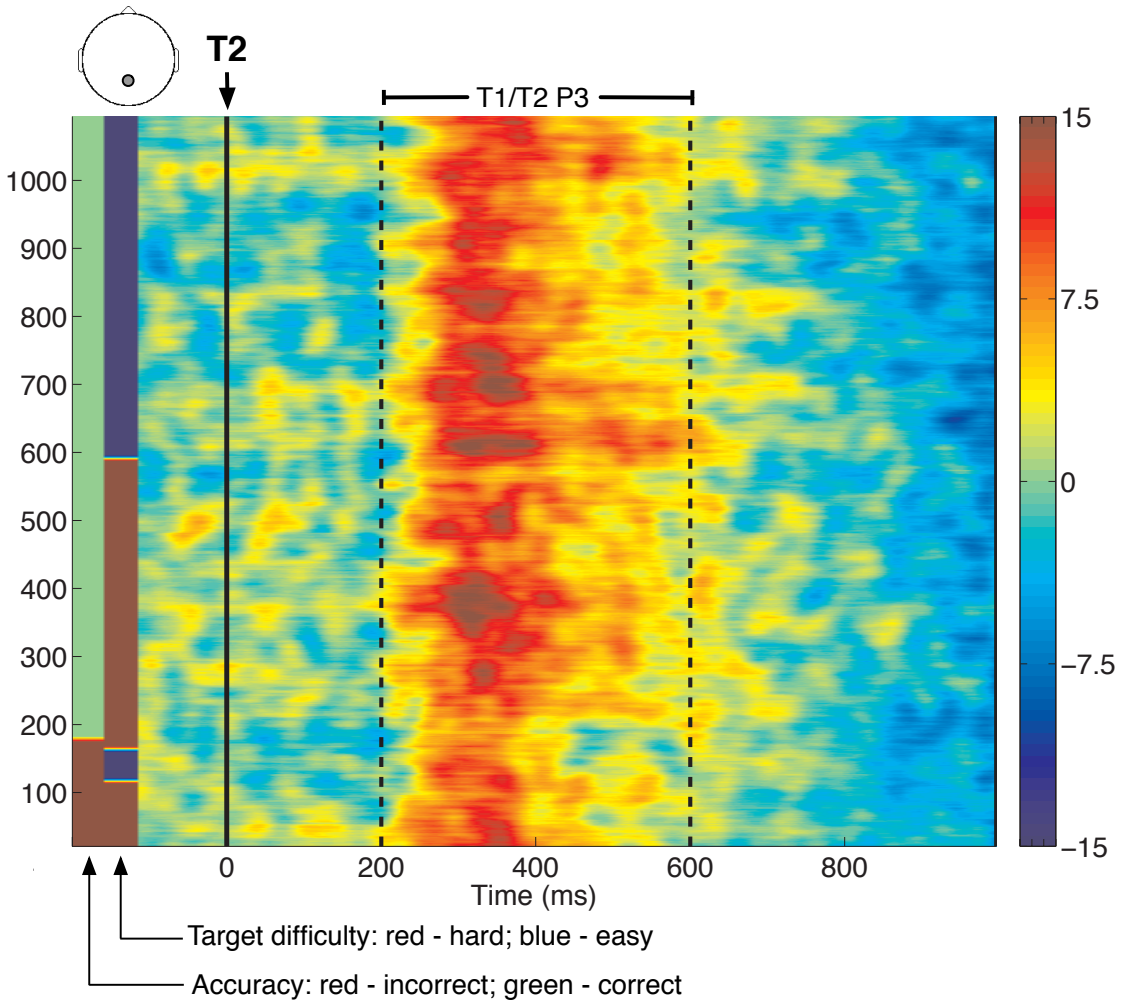


*Figure 55* ERPimage from electrode Pz for a T2 following T1 at lag 1. Colour strips to the left of the plot indicate the accuracy and target difficulty for that particular trial. Trials within each accuracy/target difficulty category are randomly shuffled before vertically smoothing to average over individual subject differences.

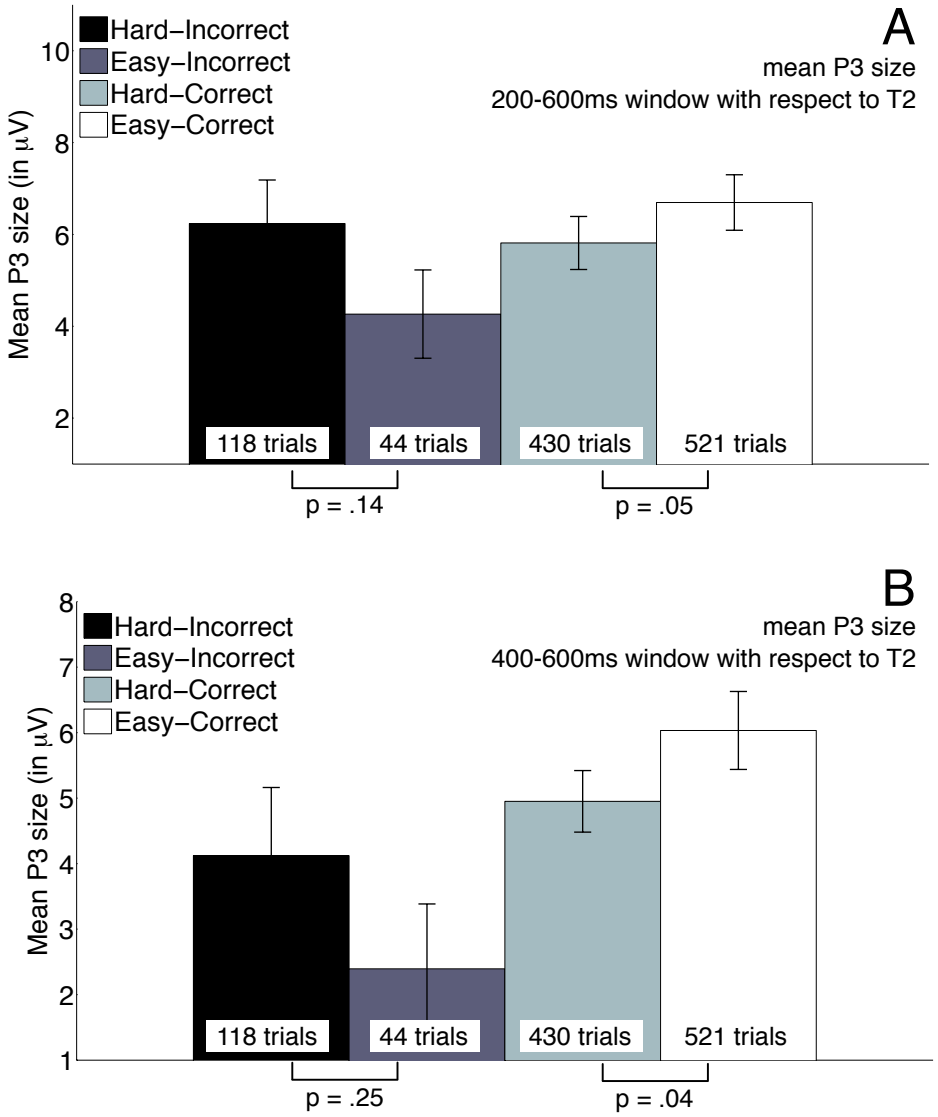


*Figure 56* Bar chart displaying the mean P3 size for each accuracy-target difficulty combination. The error bars depict the standard error of the mean. Panel A: Mean P3 size for the 200-600ms window with respect to T2 onset, which captures the whole P3 component. Panel B: Mean P3 size for the 400-600ms window. This window captures the later part of the P3 component, which is hypothesised to be influenced more strongly by T2 than by T1.

be confounded by T1, which is also influencing the P3. Nevertheless, we show how by using two different analysis techniques, the easy-hard effect on the P3 can, at least partly, validate our prediction about the EEG results for T2 presented at lag 1.

In a first analysis (depicted in Figure 56A), we employ a conservative approach and use the 200-600ms window with respect to T2 presentation to extract the mean P3 size. The 200-600ms window includes the whole P3 component for both T1 and T2. Figure 56A illustrates how for correctly reported targets, mean P3 size is larger for the Easy-Correct (6.7$\mu$V, SEM 0.6) than for the Hard-Correct (5.8$\mu$V, SEM 0.6) condition. The difference is marginally significant ($F(1,17) = 4.5$, MSE = 1.2, $p = 0.053$). Incorrectly reported T2 should not contribute towards the joint P3 component (see Kranczioch et al. (2007)), which is supported by the lack of a significant difference in P3 sizes between the Hard-Incorrect and Easy-Incorrect condition (Easy-Incorrect: 4.3$\mu$V, SEM 1.0) vs. Hard-Incorrect (6.2$\mu$V, SEM 1.0; $F(1,17) = 2.5$, MSE = 11.0, $p = 0.139$).

The second analysis (shown in Figure 56B) focusses on the later part of the joint P3 component, i.e. 400-600ms with respect to T2 presentation. This part of the joint P3 component is most likely to be influenced more strongly by T2 than by T1. As seen in Figure 56B, mean P3 size in the 400-600ms window for correctly reported targets is 6.0$\mu$V (SEM 0.6) for Easy-Correct compared to 5.0$\mu$V (SEM 0.5) in the Hard-Correct condition. The difference is more significant ($F(1,17) = 5.5$, MSE = 1.5, $p = 0.036$) than in the first analysis. This seems to support the hypothesis that the later part of the P3 is more strongly influenced by T2 and that there is indeed an 'easy-hard effect' of T2 on this part of the P3. For incorrectly reported targets, the difference becomes smaller and less significant (Hard-Incorrect: 4.1$\mu$V, SEM 1.0 vs. Easy-Incorrect: 2.4$\mu$V, SEM 1.0; $F(1,17) = 1.4$, MSE = 14.6, $p = 0.253$). As the statistical analysis becomes less significant compared to the first analysis, this further suggests that incorrectly reported T2s are not contributing towards the joint P3 component and that there is thus no difference in P3 size between the Hard-Incorrect and Easy-Incorrect conditions. Furthermore, the difference between the Easy-Incorrect and Hard-Correct conditions becomes more significant ($F(1,17) = 6.1$, MSE = 7.5, $p = 0.028$) when compared to the 200-600ms time window[6]. We would indeed expect

[6]For the 200-600ms time window, the difference between Easy-Incorrect and Hard-Correct targets only approaches significance: $F(1,17) = 3.3$, MSE = 5.1, $p = 0.094$.

there to be a significant difference in P3 size between correctly and incorrectly reported T2s at lag 1, hence, this further suggests that the 400-600ms window might be more sensitive to T2-related processing than the time window encompassing the whole joint P3 for T1 and T2 at lag 1.

**Discussion** The two-phase theory provides us with a prediction concerning the effect of 'easy-hard' on behavioural accuracy scores and P3 sizes for T2s presented at lag 1. First, we find a significant difference in accuracy scores between easy and hard target letters. Hence, the prediction concerning the behavioural accuracy scores is validated by the experimental data. Second, we find a significant difference in P3 size between easy and hard targets that are correctly reported. The profile of P3 effects for T2 presented at lag 1 is more like a target outside the AB than a target inside the AB, which is in line with the prediction from the two-phase theory.

However, the P3 analysis performed here has a potential confound. In Chapter 6, we extensively discussed how the $ST^2$ model proposes that, at lag 1, there is competition between T1 and T2 during joint working memory encoding. This hypothesis is supported by the fact that there is only one P3 component for T1 and T2, in addition to behavioural trade-off effects at lag 1 (see Chapter 6 for further details). Hence, we do not know if the P3 effects presented here are really due to the intrinsic stimulus characteristics of T2 or if they are (at least partly) due to T1 affecting the joint P3 component.

To partially address this confound, we extend our analysis to include only the later part of the joint P3, which is assumed to be influenced more strongly by T2. We find that this second analysis increases the significance of the 'easy-hard effect' for correct targets, while emphasising that there is no statistically reliable difference between easy and hard targets when they are incorrectly reported. This suggests that the later part of the joint P3 at lag 1 is indeed influenced more strongly by T2 and that there is an 'easy-hard effect' on the P3 for T2s at lag 1. Nevertheless, the aforementioned confound does remain to some extent and, consequently, the lag 1 P3 results should be interpreted with caution.

Finally, the fact that an analysis employing the later part of the P3 instead of the whole P3 window, produces a somewhat stronger statistical effect is relevant for explaining how various factors influence the profile of the P3 component. Does the P3 reflect the working

memory encoding process as such or is it a correlate of the system encoding individual targets? As discussed in Chapter 6, a larger P3 during joint consolidation of T1 and T2 at lag 1 compared to the P3 evoked by T1 with T2 at lag 8 suggests that each target is contributing to the size of the P3. Furthermore, although the P3 at lag 1 is slightly broader than the P3 for an individual target (T1 with T2 at lag 8), it clearly does not consist of two separate P3 components that overlap in time. Consequently, the results from Chapter 6 suggest that the P3 at lag 1 reflects both targets being encoded in a single episode and that the bottom-up strength of each of the targets is contributing towards the joint P3. The results from this chapter extend this argument by suggesting that, in fact, T1 is the main contributor to the earlier part (200-400ms) of the joint P3, whereas T2 contributes more significantly to the later part (400-600ms) of the joint P3.

## 8.5 Conclusion

This chapter investigates the influence of target strength and the availability of attention on target perception in RSVP. We find a difference in behavioural accuracy scores between target letters belonging to the easy and hard target categories (as defined in Chapter 6) both for targets outside the AB and also for targets inside the AB. This analysis is extended to the P3 component. For targets outside the AB, the P3 component is affected by target difficulty, whereas there seems to be no effect of target difficulty on the P3 for targets presented inside the AB.

We show that the $ST^2$ model cannot account for these findings. As illustrated by the virtual P3 component, the $ST^2$ model predicts that the P3 component should be affected by target strength (its analogue of target difficulty in the human data), no matter if the target is presented outside or inside the AB. We propose a modified theory that explains the findings of this chapter by proposing two phases of target perception in RSVP. The first phase is strongly sensitive to target strength and determines whether a target can be behaviourally reported, whereas the second phase is weakly sensitive to target strength, but influences the profile of the P3 component. We show how the timepoint of blaster enhancement modifies the profile of target activation and accounts for both the results of this chapter and also previous findings reported by Dehaene and colleagues (Sergent &

Dehaene, 2004; Sergent et al., 2005; Del Cul et al., 2007). Finally, we use the two-phase theory to make a prediction about the effect of target difficulty on a T2 presented at lag 1 and (at least partly) validate this prediction using experimental data.

This chapter provides further data investigating how targets are processed in RSVP, and, in particular, how target processing is modulated by the AB. Furthermore, we use the virtual ERP technique to validate the $ST^2$ model against the experimental data. Unlike previous chapters, however, the data presented here is in contrast with the $ST^2$ model. In response to this discrepancy, we show how we can borrow concepts from the $ST^2$ model to propose a modified theory that provides an explanation for the data presented in this chapter.

# Part IV

# Further discussion & conclusion

# Chapter 9

# Using EEG to design adaptive computer systems

## 9.1 Introduction

Research into visual attention is highly relevant to the design of computer systems, as the human interface to a system is often a major processing bottleneck. This chapter (based on a technical report published as Wyble et al. (2006)) contains a proposal of how some of the theoretical work presented in this thesis could be applied to the practical design of an adaptive computer system.

Findings from RSVP studies and also the AB paradigm provide insights into the temporal limits of human attention. These should be considered when designing computer systems in order to avoid presenting messages and signals while the user's attention is 'blinking' (Su, Bowman, & Barnard, 2007). However, although Human-Computer Interface (HCI) design is constantly being optimised to increase the inherent salience of signals directed at the human user (using warning lights and sounds, for example), accidents still occur because human users fail to perceive critical warning signals.

If we take computer networking as a metaphor, the interaction between a human user and a computer system is similar to transmitting information over a communication channel without acknowledgement of receipt. The computer is transmitting information to the user with the intent that it be received, but the computer system has no way of verifying that

this transmission was successfully completed. The goal of the approach proposed in this chapter is for the computer system to receive such an acknowledgement signal from the user using EEG. Previous work in the field of brain-computer interfaces (BCI) has mainly focused on *human-to-computer* interaction (e.g. Donchin, Spencer, & Wijesinghe, 2000; Schalk, McFarland, Hinterberger, Birbaumer, & Wolpaw, 2004; Blankertz et al., 2006). Our approach uses EEG to enhance the reliability of *computer-to-human* interaction, an application that is virtually unexplored. We describe a system that uses EEG to inform a computer that its user may have missed a critical piece of information. This warning will allow the computer to re-present missed information to the user until the message has been perceived.

Our approach proposes a device with the following characteristics:

- Small enough to be enclosed within a helmet
- Easily shielded from nearby interference
- Sufficiently low power consumption to run on lightweight battery power
- Extremely rapid response (response within less than 1 second of target onset)
- Easy to set up

These restrictions rule out both sophisticated waveform analyses and multi-electrode arrays. The focus is thus on whether a relatively simple analysis of data recorded from a single pair of electrodes can provide useful information to the computer about whether a target was detected or not.

## 9.2 Method

The analysis uses EEG data from Experiment 1, as described in the appendix. We rely on two elements of the EEG signal as indicators of target perception:

- The P3 component of the ERP (see Section 2.1.1), which is often distinctive enough to be detected on a trial-by-trial basis, i.e. in the raw EEG signal *(raw P3)*. A measure of total area under the curve - centred around the time of maximal P3 amplitude - is computed for each participant. This time window ranges, at most, from 300-700ms after target presentation, but varies from subject to subject. Both trials in which the target is correctly reported (*target-seen* trials) and trials where the target is incorrectly reported (*target-missed*

trials) are included in the analysis. For each participant, a specified percentage of the area under the P3 component (within the subject-specific time range) from the grand average ERP of all target-seen trials is taken as threshold. Then, for each trial, we determine if the raw P3 of that trial exceeds the threshold. If a target-seen trial has a raw P3 larger than the threshold, the trial is counted as a hit, a raw P3 below threshold means that the trial is scored as a miss. For target-missed trials, a raw P3 above threshold is counted as a false alarm, a raw P3 below threshold categorises that trial as a correct rejection.

- Changes in the power of EEG oscillations near 10 Hz are known as the alpha power band (see Section 2.1.2). Alpha band oscillations are often clearly visible in the raw EEG signal. In the second analysis, we compute the *raw alpha power* using a fast fourier transform (FFT) in the range of 8-12 Hz for a time window of 400-900ms after target presentation. A threshold for each subject is determined by taking a specified percentage of the subject's alpha power averaged over all target-seen trials. Alpha power tends to diminish with the onset of a cognitive event, such as detecting a target, encoding something into memory or initiating a movement (Pfurtscheller & Lopes da Silva, 1999). Consequently, if on a target-seen trial raw alpha power is below threshold, this trial is counted as a hit, a trial above threshold means that the trial is scored as a miss. For target-missed trials, raw alpha power below threshold is counted as a false alarm, raw alpha power above threshold categorises that trial as a correct rejection.

With respect to our proposed adaptive computer system, a hit means that the system registers that a target has been perceived by the user and, thus, this target does not have to be presented again. A false alarm means that the system 'thought' the target was perceived even though the user did not see it. Consequently, the algorithm needs to be optimised in order to maximise hits and minimise the number of false alarms.

## 9.3 Summary of results

The following section contains a summary of the results. Please refer to Wyble et al. (2006) for a full description of the results.

### 9.3.1 Separate P3 and alpha analysis

We begin by separately testing each of our algorithms. We first use the raw P3 and then raw alpha power as a method for target detection. In the P3 analysis, we use 50% of the area under the raw P3 curve as the threshold for the algorithm. For one subject, where the raw P3 analysis works particularly well, hits are as high as 62% with only 8% false alarms. When averaging across all subjects, however, this method produces a relatively high number of hits (60%), but the rate of false alarms is also rather high (40%). Such a high number of false alarms is troublesome for a safety-critical system, because the system falsely assumes that a warning message has been perceived by the user, which could have serious consequences.

The analysis using raw alpha power performs worse than the raw P3 analysis. When we use a threshold of 150% of a subject's alpha power average, we get a very high percentage of hits (91% when averaged across all subjects), however, this is accompanied by an unacceptably high number of false alarms (83% when averaged across all subjects).

### 9.3.2 Combined P3 and alpha analysis

| Subject | Hit | False alarm |
|---|---|---|
| 1 | 28 | 14 |
| 2 | 33 | 10 |
| 3 | 33 | 10 |
| 4 | 29 | 3 |
| 5 | 25 | 8 |
| 6 | 29 | 12 |
| 7 | 33 | 12 |
| 8 | 24 | 4 |
| 9 | 25 | 5 |
| 10 | 25 | 16 |
| 11 | 31 | 27 |
| 12 | 29 | 12 |
| Average | 29 | 11 |
| Standard error | 0.97 | 1.85 |

*Table 2* Results from the combined P3 and alpha detection algorithm.

A third analysis combines both previous attempts (see Table 2). With both raw P3 and raw alpha power as detection criteria, it is more difficult for a target to be classified

as either a hit or a false alarm, as each trial now has to pass two tests. Furthermore, we adapt the thresholds used in the algorithms of P3 and alpha power tests, in that raw P3 area has to exceed 90% (rather than 50%) of the average P3 area and raw alpha power has to be less than 100% of the average alpha power (rather than 150%). This combined approach improves the algorithms performance compared to using solely the raw P3 or raw alpha power as detection criteria. The numbers of false alarms are drastically reduced (average false alarms 11%), which is an important improvement if this type of algorithm was employed in a safety-critical environment. However, the reduction in false alarms comes at the expense of decreasing the number of hits (average hits 29%). Results vary across subjects, but for six of the subjects, false alarms are at or below 10%, with hit rates above or around 25%.

## 9.4 The practical details of such a device

In order to ensure that a critical piece of information is perceived by the user, computer systems often display a salient visual or auditory alarm signal to capture the user's attention. However, if an environment is particularly crowded with such salient signals, the user can be faced with an overwhelming number of such signals, which means that some signals will be ignored. The Brainwave Based Receipt Acknowledgement device (BBRA, Figure 57) uses EEG waves recorded from the user's scalp to provide the computer system with feedback about whether the user perceived a particular piece of information or not. In theory, such a device could prevent the overuse of alarm signals and the resulting information overload in information-rich environments. The system simply re-displays a critical message until it has been successfully perceived by the user, as indicated by an acknowledgement receipt from the BBRA. In time-critical environments (such as, for instance, the cockpit of a jet plane), the BBRA device would have to operate within very short time scales, hence, the acknowledgement must be sent almost immediately after the BBRA algorithm has registered successful target detection. In the following, we describe how this system could be implemented.

The BBRA system is intended to fit within a head-mounted device (Figure 57A). It

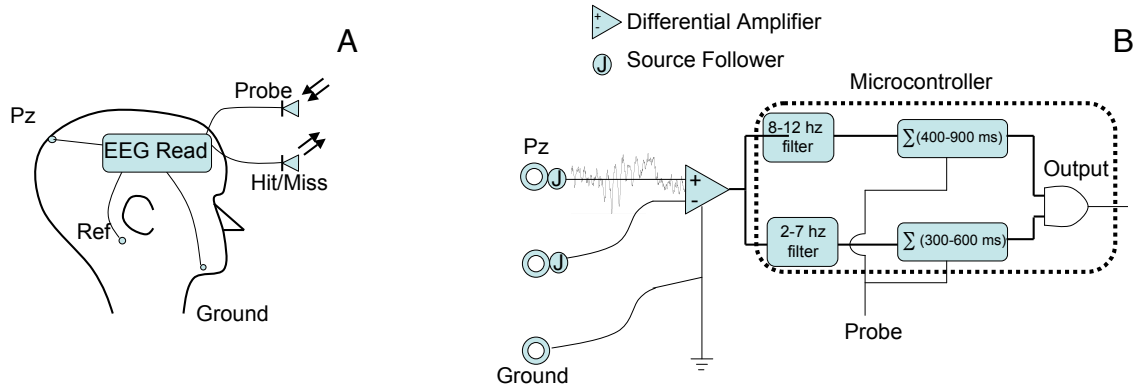


*Figure 57* Panel A: Illustration of the proposed BBRA system. The probe input provides the time lock signal indicating target presentation. The 'hit/miss' output informs the computer system whether the target was seen by the user or not. Panel B: Schematic diagram showing the electrodes and microcontroller circuit for the BBRA.

requires three electrodes held against the scalp with an elastic headband and a small circuit-board with components powered by a battery pack. The EEG signal needs to be amplified, which could be done using 'active electrodes'[1], which allow an EEG signal to be recorded in noisy environments without too much interference from motion artifacts. The EEG would be recorded by taking the difference between voltages recorded at two positions on the scalp (Figure 57A). Firstly, we require the Pz electrode location (suitable for P3 and alpha oscillatory analysis) at the back of the head over the parietal cortex. Second, we need a reference electrode situated somewhere on the scalp but as far away as possible from Pz. Finally, we require the ground channel, which can be placed anywhere convenient. The Pz and the reference electrode are both amplified using active electrodes, while the ground channel is a direct connection to the ground rail of the amplifier circuit. The two electrode signals are fed into a differential amplifier, which produces a single output voltage that will be used by the following parts of the circuit.

The EEG waveform is passed through two parallel circuits, one for each of the detection mechanisms (Figure 57B). The signal intended for alpha detection begins with a filter that selects frequencies in the 8-12 Hz range, which are then rectified (the equivalent of taking the absolute value of the signal). The amplitude of the resulting signal, when summed over the specified time window, reflects the amount of power in the alpha range. The other

[1] see e.g. `http://www.biosemi.com/active_electrode.htm`

circuit is filtered from 2-7 Hz, which isolates slower ERP components, such as the P3. The raw P3 wave can then be integrated over the specified temporal window. In the next step, the recorded signal is passed into a microcontroller, which converts the analogue waveform into a digital signal.

After receiving the time locking signal of target presentation (e.g. via an infrared input - called 'Probe' in Figure 57B), the microcontroller takes the sum of both (now digitised) input channels over the specified range of time points. If these exceed the corresponding thresholds for both the alpha and P3 detection algorithms, the BBRA registers that the target message was seen by the human user and informs the computer system accordingly via the output signal (i.e. the 'Hit/Miss' infrared output signal shown in Figure 57B). If the threshold is missed, however, the BBRA sends a negative feedback via the output signal and the computer system is prompted to present the target message again.

## 9.5 Evaluating the approach

The work presented in Wyble et al. (2006) is a preliminary exploration of using EEG signals to enhance the reliability of the interaction between a computer system and its human user. It shows how insights from EEG studies investigating temporal attention can aid the development of adaptive computer systems.

Although the performance of the algorithm is far from perfect, we believe the approach makes a contribution due to its simplicity, which makes it feasible to run this algorithm on a lightweight computer system. Furthermore, we expect the results from a real working prototype of this system to be superior to the results reported here in a number of respects. First, an actual head-mounted EEG recording setup with amplification applied locally to each electrode will produce more robust data than the conventional recording setup that was used to obtain the data of Experiment 1 (see the methods section in the appendix for details). Second, the participants in this study were not selected, but volunteers from the university population. Some participants' EEG data is more suitable for this kind of analysis, as is evident from the large variation in the numbers of hits and false alarms between participants. Hence, research on a BBRA system might progress more effectively on selected subjects that make ideal candidates for this sort of interface.

# Chapter 10

# Summary & conclusion

## 10.1 Summary

In this thesis, we present a novel approach of relating neural modelling of cognition to the recording of electrophysiological brain activity using EEG. We use an existing neural network model (the $ST^2$ model), which has proven itself a valuable tool for investigating and describing the mechanisms underlying temporal attention and working memory in humans (Bowman & Wyble, 2007). The $ST^2$ model, as it was originally published, is capable of replicating a wide range of behavioural data related to the attentional blink (AB). The contribution of this thesis consists of extending the $ST^2$ model to generate 'artificial electrophysiological traces' - so-called *virtual ERPs*. The virtual ERPs are used to make predictions and propose explanations for the experimental results gained from two EEG studies that were conducted to investigate target processing under high temporal demands. We now summarise the work that has been described thus far and introduce the final chapter of this thesis.

Part I contained a general introduction and was concerned with reviewing the literature that is relevant to the work presented in this thesis. In Chapter 1, we commenced with a general introduction of why it is important to understand the mechanisms that underlie temporal attention in humans, for example when designing safety-critical computer systems. Furthermore, Chapter 1 elaborated on how predictions generated from formal cognitive models can help the design of electrophysiological experiments while, at the same time,

providing a powerful constraint for validating the neural models themselves. Chapter 2 contained a review of the relevant literature. We described how the EEG technique is used to investigate cognitive processes by analysing how brain signals recorded non-invasively from the participant's scalp are modulated in response to a cognitive event. Following this, we gave a brief introduction into cognitive modelling techniques and focussed on those models that employ computational techniques, such as artificial neural networks. Chapter 2 concluded with an overview of the literature on visual processing under high temporal demands. We began by introducing the rapid serial visual presentation (RSVP) paradigm, where items are presented at a very rapid rate in the same spatial location. As each item masks its predecessor, item representations in RSVP can be weakened to an extent where people are unable to detect the item's identity, even though such an item would be easily perceived if it was presented individually. The RSVP paradigm is a tool for studying the AB, which seems to provide insights into the temporal limitations of conscious perception. In the following, we provided an overview of the AB phenomenon and reviewed previous experimental studies, which have investigated the AB using behavioural, electrophysiological and brain imaging techniques.

Following the introduction, Part II was concerned with describing current theories of the AB and presenting our novel approach of generating virtual ERPs from the $ST^2$ model. Chapter 3 commenced with a description of informal theories of the AB. We presented the textual models that are most influential in the field of AB-related research. Following this, we described current formal models of the AB and focussed on those models that employ neural network architectures. Chapter 3 concluded with an extensive description of the $ST^2$ model. This neural network model provides the basis for our novel approach of generating virtual ERPs, which was described in Chapter 4.

Part III presented the experimental results of the two EEG studies that were conducted to investigate target processing under high temporal demands and, in particular, during the AB. Chapters 5, 6, 7 and 8 presented detailed investigations with respect to different theoretical issues relevant to the study of temporal attention. How each of these chapters contributes towards the understanding of target processing under high temporal demands and, in particular, theories of the AB is discussed in Section 10.2 of the current chapter. Furthermore, the chapters in Part III showed how the virtual ERP technique can be used to

make predictions and propose explanations for the experimental results. In other words, the work presented in Part III validated the virtual ERP method as a useful tool for scientific investigation.

This thesis concludes with Part IV. Chapter 9 contained a proposal of how the design of computer systems might benefit from research on temporal attention, such as the findings presented in this thesis. Chapter 10 began with a summary of the work presented so far and continues with a section on the contributions that this thesis has made. Finally, Chapter 10 ends by considering how the work presented in this thesis could be extended and further developed in future investigations.

## 10.2 Contributions

This thesis has made a number of contributions to understanding the mechanisms underlying temporal attention in humans. Furthermore, we have proposed a technique for using a neural model to generate virtual ERPs. This section begins with a discussion of the implications for theories of target processing in RSVP and, in particular, the AB. After this, we elaborate on how the virtual ERP technique can benefit neural modelling of cognition and the design of EEG experiments.

### 10.2.1 The influence of distractors on target processing

Chapter 5 compares the EEG profiles for targets presented individually (followed by a single mask in so-called skeletal presentation) to targets presented within a continuous RSVP stream of distractors. We find differences in the EEG correlates of target processing all the way from early visual processing to the encoding of targets into working memory. First, skeletal targets evoke a relatively large neural response at occipital areas. For targets in RSVP, however, target-related activation is 'hidden' within the continuous ssVEP wave caused by the items occurring in the RSVP stream.

Furthermore, we find that the presence of distractors strongly modulates the way in which targets are selected. The fact that a target is surrounded by distractors when presented in an RSVP stream seems to enforce a strategy of late selection. In other words, observers have to process items rather extensively before they can distinguish targets from

surrounding distractors. We hypothesise that this difference in selection strategy is reflected by a modulation of both the P2 and P3 ERP components, as suggested by our EEG results. We then go on to modify the $ST^2$ model to simulate skeletal presentation. We find that by making minor changes to the architecture of the model, we can achieve a qualitative match of the behavioural data. Furthermore, we qualitatively replicate the EEG effects using virtual ERPs. This demonstrates that - subject to minor changes - the $ST^2$ model can be adapted to simulate the experimental effects in related paradigms.

### 10.2.2 The meaning of P3 amplitude in RSVP

The P3 component is one of the most studied aspects in ERP research. Nevertheless, the neural mechanisms underlying the P3 remain the subject of much debate. One controversy concerns the meaning of P3 amplitude for targets in RSVP. It is often argued that the amplitude of the P3 reflects the amount of 'cognitive resource' invested in the task (e.g. Shapiro et al., 2006). Indeed, experimental studies have confirmed that when subjects are required to invest more effort in a task and, critically, subjects know this *beforehand*, P3 amplitude increases (Sirevaag et al., 1989; Wickens et al., 1983; Kramer & Hahn, 1995). In Chapter 5 we find that for the skeletal task, which is behaviourally easier than the task of detecting targets in RSVP, P3 amplitude decreases. Hence, our results from Chapter 5 support the hypothesis that the P3 component amplitude is a correlate of the amount of effort invested in the task, if subjects can allocate this effort beforehand.

In Chapter 6, however, we show that the P3 cannot always serve as an index of the amount of resource allocated to the task. Rather, it depends on the circumstances of the experimental paradigm. If the target is presented at random in an RSVP stream, subjects cannot predict when the target will occur. Furthermore, subjects do not know what the identity of the target will be on a particular trial and whether that target will be easier or harder to detect. Hence, although there might be random fluctuations from trial to trial, on average, pre-allocated effort should be equal on each trial. Our results, however, show a larger P3 wave for those target letters with higher accuracy scores (easy letters). The 'easyness' of these target letters is solely due to intrinsic stimulus characteristics. According to the 'resource-hypothesis', however, subjects should require less effort to detect easy letters and consequently have a smaller P3 for easy letters. Hence, the EEG results from Chapter 6

provide evidence for the hypothesis that - if pre-allocated effort is equal (as is the case for targets in RSVP) - P3 amplitude is not an index of the amount of resource invested in the task, but rather the P3 is modulated by the amount of bottom-up strength of the target (see also Kok, 2001).

### 10.2.3 Implications for theories of the attentional blink

The results presented in this thesis have a number of implications for theories of the AB. We discuss the implications first for the two-stage theory (Chun & Potter, 1995) and the related $ST^2$ model (Bowman & Wyble, 2007). Second, we elaborate on our results' implications for the two-stage theory's main competitor, the interference or resource sharing theory (Shapiro, Raymond, & Arnell, 1997; Shapiro et al., 2006).

#### Two-stage theories

As discussed in Section 3.1.2, two-stage theories emphasise three important mechanisms of the visual system. The parallel first stage performs early visual processing and has no capacity limitation. The serial second stage, which reflects working memory encoding, has limited processing capacity as it encodes items sequentially. Third, a transient attentional enhancement mechanism enhances target representations, which assists the working memory encoding process.

The EEG results from Chapter 6 support a notion of serial working memory encoding during the AB. The findings are thus entirely consistent with the predictions from the two-stage theory, which is illustrated by the close match between the human and virtual ERPs. At lag 1, however, the EEG results suggest that working memory encoding is not serial, instead, the targets are encoded into working memory together. In addition, the behavioural results suggest a trade-off in accuracy between the two targets at lag 1. Two-stage theories account for these results, because they argue that - at lag 1 - the targets are encoded together in a single episode (see also Wyble et al., 2009). As there is competition between targets during joint working memory encoding, this accounts for the observed trade-off effects in accuracy scores.

Chapter 7 provides evidence for increased temporal variance in target processing for targets presented during the AB when compared to targets that are presented individually.

This finding is accounted for by the serial nature of working memory, which is inherent to two-stage theories. During the AB, the second target (T2) cannot be consolidated until the encoding of the first target (T1) has completed. As the duration of T1's encoding process depends on its strength value, the time point of T2 processing will vary accordingly.

In general, the two-stage theory can account for the EEG results from Chapter 8. Outside the AB, attention is readily available and target perception depends on the bottom-up strength of the target. Hence, the P3 is strongly influenced by target strength, as seen in the results of Chapter 8. According to two-stage theories, the serial nature of working memory encoding during the AB is enforced because attention is suppressed while T1 is encoded, which ensures that T2 does not corrupt T1's working memory encoding process. Consequently, whether attention is available or not becomes the main determinant of target perception during the AB. The $ST^2$ model cannot account for the EEG findings for targets during the AB, because it suggests an influence of target strength on the P3 component when targets are presented during the AB. However, the two-phase strength sensitivity theory described in Chapter 8 (which is also based on the two-stage theory) provides an explanation for both the behavioural and EEG results for targets outside and inside the AB.

#### Interference or resource sharing theory

In contrast and as described in Section 3.1.3, the interference or resource sharing theory suggests that there is competition between items during working memory encoding and that this competition can last throughout the duration of the AB. Such a theory thus argues that there is mutual interference during the AB, in that both targets influence the other's processing.

The EEG results from Chapter 6 provide evidence against resource sharing or interference during the AB. We find that T2 processing is influenced by T1 and, indeed, this is reflected in the impairment of T2 accuracy (i.e. the AB). On the contrary, however, T1 processing is not influenced by whether T2 is presented inside or outside the AB. In addition, T1 processing is unaffected no matter if T2 is successfully detected or missed during the AB. Hence, the behavioural and EEG results from Chapter 6 argue that there is unidirectional influence during the AB (i.e. T1 *does* influence T2) but the interference is

not mutual (i.e. T2 *does not* influence T1). Mutual interference only obtains at lag 1, where trade-off effects and joint consolidation suggest competition between targets. Consequently, the theoretical arguments underlying the interference/resource sharing theory are limited to such short timespans, i.e. if T2 is presented at lag 1. If T2 is presented during the AB, however, T2 does not share resources with T1 nor does T2 compete with the preceding T1 (see also Bowman & Wyble, 2007; Craston et al., 2009; Wyble et al., 2009).

As previously mentioned, the resource sharing or interference theory argues that T2 competes with T1 during the AB. T2 strength varies from trial to trial, so, if there was indeed competition between T1 and T2 during the AB, T2's impact on T1 should also vary from trial to trial. Consequently, on some trials, T2 will be able to win the competition at an earlier time point than others, which would lead to the temporal variance in processing of correctly reported T2s, as observed in Chapter 7. This line of argument, however, suggests the same should be true for T1 processing. As with T2, T1's strength also varies from trial to trial, hence, according to interference/resource sharing models there should also be considerable amounts of temporal variance in the processing of T1. Our EEG data from Chapter 7, however, suggests that this is not the case. As quantitatively shown by the ITC analysis, there is significantly more variance in the processing of T2 as compared to T1. Hence, the EEG results from Chapter 7 cannot be fully accounted for by the interference/resource sharing theory.

Finally, the EEG results from Chapter 8 suggest that conscious perception (as reflected by the P3 component) is mainly influenced by target strength for targets presented individually (i.e. outside the AB). This first result from Chapter 8 is consistent with the interference/resource sharing theory. A stronger target is more resilient to interference from distractors in visual short term memory and is thus more likely to be successfully encoded into working memory. The second EEG result from Chapter 8 concerning targets during the AB, however, is troublesome for the interference/resource sharing theory. If there was indeed competition between T1 and T2 during the AB, conscious perception during the AB should be strongly modulated by T2 strength. After all, strong T2s should be able to win the competition against T1 and surrounding distractors during the AB, which would cause them to be consciously perceived. Our results suggest, however, that target strength is only a minor factor influencing conscious perception during the AB, hence, this

is in contrast with the interference or resource sharing theory.

### 10.2.4 Virtual ERPs as an additional dimension of neural modelling

Assessing the quality of computational models is not straightforward, especially when trying to pick 'the best model' amongst a group of competitors (Pitt & Myung, 2002; Penny, Stephan, Mechelli, & Friston, 2004). Sophisticated assessment techniques indicate that the closeness of fit with experimental data is important (Massaro, Cohen, Campbell, & Rodriguez, 2001) but cannot be the sole criterion for evaluating a computational model (Pitt, Myung, & Zhang, 2002; Pitt, Kim, & Myung, 2003). Indeed, Roberts and Pashler (2000) suggest that the best strategy for testing 'a theory with free parameters is to determine how the theory constrains possible outcomes (i.e. what it predicts), assess how firmly actual outcomes agree with those constraints, and determine if plausible alternative outcomes would have been inconsistent with the theory'. Thus, the more possibilities there are to make predictions and impose constraints on the model, the better (see also Popper (1959)).

If a model generates virtual ERPs as well as simulating behavioural output, this provides the means to make predictions in the domain of electrophysiology, which improves the usefulness of the model as a tool for experimental research and theoretical reasoning. It also increases the number of constraints imposed on the model. In addition to having to replicate a 'static' set of behavioural data, the model's neural activation dynamics become important as they have to follow the profile observed in the human ERP for the virtual ERP to be a good match of its human counterpart.

Furthermore, increasing computing power allows neural models to become more and more complex. In these complex systems, it can become difficult to comprehend exactly how individual nodes of the model are behaving over time. Thus, when building and using a computational model, it is important to be able to visualise the behaviour of the model. Grand average virtual ERPs illustrate activation profiles over time, however, an even more detailed analysis of the activation dynamics for a particular parameter setting is possible when using virtual ERPimages. In this thesis, we have shown how virtual ERPimages depict the neural activation profiles of individual trials of a simulation run and, hence, provide a comprehensive illustration of the behaviour of the neural model.

### 10.2.5 Virtual ERPs assist electrophysiological experimentation

The virtual ERP technique also provides opportunities for electrophysiological experimentation strategies. A review by Picton et al. (2000) emphasises the importance of a clear hypothesis before conducting EEG experiments: 'The overwhelming amount of ERP data along the time and scalp-distribution dimensions can easily lead to incorrect post hoc conclusions based on trial-and-error analyses of multiple time epochs and electrode sites.' Unlike the predictions that can be derived from textual theories, virtual ERPs from neural models provide a means of making more formal predictions of ERP latencies and amplitudes, which can aid the construction of hypotheses prior to experimental design and data collection. One can investigate how parameter changes in the model affect results in both the simulated behavioural and virtual 'electrophysiological' domain, thereby giving a principled method for exploring a theoretical hypothesis.

## 10.3 Future work

This section concludes this thesis by discussing possible extensions to our work. In Chapters 6, 5, 7 and 8, the virtual ERP approach has proven itself a powerful tool for making predictions about experimental data and validating the $ST^2$ model. Amongst other issues, the final section discusses a potential method for modelling individual differences during the AB and elaborates on how a tighter connection of the $ST^2$ model with brain anatomy would open up a range of possibilities for modelling data from, for instance, studies on patients with brain lesions and also fMRI experiments.

### 10.3.1 Additional evidence elucidating the nature of working memory encoding during the AB

Chapter 6 provides evidence in favour of serial working memory during the AB, i.e. suggesting that working memory encoding of the second target is delayed until the first target has been successfully encoded. Our results thus provide evidence against the resource sharing theory of the AB.

However, why is it that a number of previous studies reported their EEG/MEG results to be evidence in favour of resource sharing during the AB? First, we discuss in Section 6.5.4

why we believe that these previous studies fail to provide definitive proof of resource sharing during the AB. Second, these findings were interpreted under the assumption that the P3 provides a measure of how much processing resource was allocated to the target. As discussed in Chapter 6, such an assumption does not hold for targets in RSVP, rather, P3 size is correlated with the amount of bottom-up strength of that target.

Nevertheless, the variance in the results between the studies is intriguing. If the AB does indeed reflect serial working memory encoding, all EEG data should be similar to the results presented in Chapter 6, no matter if the data is presented as part of an article arguing *in favour of* or *against* resource sharing during the AB.

Another example for this inconsistency comes from a study investigating the effects of meditation on the AB (Slagter et al., 2007). The authors have two groups of subjects (one consists of normal controls and the other contains meditation practitioners) participating in an AB task while recording their EEG. Interestingly, they find that for practitioners, the T1 P3 component is larger for trials when T2 is missed during the AB and smaller if T2 is seen during the AB. Those participants thus show exactly the kind of trade-off in T1 P3 sizes that we argue does indeed provide evidence in favour of resource sharing during the AB (as defined in Section 6.5.4). In the second group of participants (normal controls/novices at meditation), however, there is no such effect. Their T1 P3 has the same size no matter if T2 is seen or missed during the AB, thus replicating our results from Chapter 6. Hence, Slagter et al. (2007) find EEG data indicating serial working memory encoding for 'normals' but EEG results supporting resource sharing for practitioners and all of this in one single study.

In consequence, further experiments would seem to be required to provide definitive evidence for serial working memory encoding during the AB. These experiments should place particular emphasis on avoiding confounding factors, such as comparing the EEG signals between different groups of participants (Slagter et al., 2007). Rather, a future study should present a comparison between the EEG traces for the AB and noAB conditions only for those participants that show a clear AB (i.e. *blinkers* according to the definition of Martens, Munneke, et al., 2006).

### 10.3.2 Dissecting the virtual ERP to identify the neural substrates of the human EEG

Due to the nature of EEG, the isolation of signals related to the cognitive processes of interest from background activity can be problematic. The virtual ERP, however, can be dissected into its underlying components. For example, one can generate virtual ERP traces related to attentional processes or working memory consolidation by including only the associated parts of the $ST^2$ model. If one used blind source separation techniques, such as independent component analysis (ICA; Makeig et al., 1999; Makeig, Debener, et al., 2004; Makeig, Delorme, et al., 2004; Debener, Makeig, Delorme, & Engel, 2005), to decompose the human ERP, correlations between individual components of the virtual ERP and the human ERP might help to further explain the cognitive processes underlying the human ERP.

Virtual ERPs may also help provide insights into investigating the neural substrates of the human EEG. The localisation of human ERPs is restricted by the inverse problem (von Helmholtz, 1853) and relies on sophisticated algorithms for source analysis, which are based on a number of assumptions and approximations (see Section 2.1.3). The origins of activity contributing towards the overall waveform of a virtual ERP, however, can be localised to certain parts of the neural network model. In order to further associate layers of the model with human ERP components, however, one has to link parts of the $ST^2$ model to specific brain areas, which was only partially done in Bowman and Wyble (2007).

Traces related to early visual processing are generated from the input and masking layer of the $ST^2$ model. We suggest that these layers are responsible for early visual processing that, in the human brain, is performed in areas of occipital cortex. This is justified by the input layer receiving 'artificial visual input', which is then passed on to the masking layer. The analogue of these layers in the brain could be visual areas in occipital cortex receiving input from the retina via the parvo- and magnocellular pathways.

The P3, on the other hand, is far more difficult to pin down. This is reflected in the way the virtual P3 is generated from the $ST^2$ model, as the virtual P3 is a summation of activity from nearly all layers beyond early visual processing. The virtual P3 thus cannot really inform a source localisation of the human P3 component. In the human ERP, the

P3 is largest at parietal and central electrodes. However, and this is especially the case if the P3 component for target items is as distinctive as it is in RSVP, it can be measured at electrodes located throughout the scalp. Nevertheless, recent combined EEG and fMRI studies have suggested that the P3 is generated in inferior parietal areas including the temporoparietal junction (Bledowski, Prvulovic, Hoechstetter, et al., 2004; Strobel et al., 2008). In Bowman and Wyble (2007), it was suggested that the blaster of the $ST^2$ model might be located in the temporoparietal junction. Hence, this association could provide a starting point for creating associations between the generators of the virtual and the human P3 component.

### 10.3.3 'Lesioning' the $ST^2$ model

Computational models can benefit experimental investigations into the effect of brain lesions on human perception. For instance, a model called the 'theory of visual attention' (TVA; Bundesen, 1990) demonstrates how a computational model can be a useful tool for making predictions about impaired behavioural accuracy scores and reaction times observed in patients with brain lesions (Habekost & Bundesen, 2003).

Intracranial but also non-invasive electrophysiological recording is an essential part of clinical neurophysiology. Although relatively poor in terms of spatial resolution, EEG is a useful tool for investigating the deficits caused by brain lesions and other neurological diseases. EEG can be applied without much discomfort to the patient and allows the researcher to record neural brain activity with excellent temporal resolution.

To date, the $ST^2$ model has not been used in patient studies. Again, and as discussed in the previous section, as part of future work one would first have to establish a tighter link between parts of the model and specific brain regions. Subsequently, certain parts of the model could be 'artificially lesioned' to test the predictions the model would make both in terms of simulated behavioural and also virtual ERP data. These predictions could then be verified using data collected from patients suffering from brain lesions.

### 10.3.4 Virtual fMRI traces?

Recently, a number of articles have presented the results of combined EEG/ERP and fMRI experiments. There is debate about the superiority of simultaneous recording of EEG and fMRI versus other EEG-fMRI studies that employ each technique in a separate session (Bledowski, Linden, & Wibral, 2007; Debener, Ullsperger, Siegel, & Engel, 2007; Strobel et al., 2008). However, simultaneous EEG and fMRI recording (e.g. Debener, Ullsperger, et al., 2007) has at least one major advantage over combined but separate EEG-fMRI experiments (e.g. Bledowski, Prvulovic, Goebel, Zanella, & Linden, 2004): it is a well-known fact that subject behaviour often fluctuates from session to session. Hence, only the simultaneous approach ensures an equal experimental setting in both the recording of the EEG and the fMRI data (Debener, Ullsperger, Siegel, & Engel, 2006).

Debener et al. (2006) have proposed an innovative approach to the simultaneous recording of EEG and fMRI. Once the EEG data has been recorded using an fMRI-compatible EEG system, fMRI-induced artifacts are removed using specialised correction algorithms (for details see Debener, Strobel, et al., 2007; Debener, Mullinger, Niazy, & Bowtell, 2008). Following this, the EEG data is unmixed using ICA, which allows the removal of other artifacts (such as eye-blinks) and often provides good results at separating the brain-related component activations from the rest of the EEG data. Then the data is analysed at the single trial level and convolved with a hemodynamic response function, which corrects for the temporal delay between the EEG and fMRI response. Finally, the EEG data is localised to brain regions using a model of equivalent dipoles and this data is used to predict the BOLD response of the fMRI signal. Consequently, Debener et al. (2006)'s approach essentially uses the EEG data as a model for predicting and validating the fMRI response and associated hypotheses. This approach has been employed successfully in various experimental studies (Debener, Ullsperger, et al., 2005; Strobel et al., 2008).

A tighter link with brain anatomy could open up possibilities of applying the $ST^2$ model's predictions and explanations to fMRI data. The dynamic causal modelling approach (Friston, Harrison, & Penny, 2003) has demonstrated the benefits of models that are capable of simulating both EEG/MEG and fMRI data (David, Kiebel, et al., 2006). Hence, it would be useful to extend the $ST^2$ model to predict fMRI activity. As such work

would obviously require a major extension to the $ST^2$ model, it remains a prospect for future work. However, if such a step was realised, it would open up the possibility for 'virtual fMRI traces' from the $ST^2$ model, which would complement the virtual ERPs. Such an approach would enable the $ST^2$ model to make a contribution to the dynamic field of combined EEG-fMRI recording.

### 10.3.5 Modelling blinkers and non-blinkers

Like other psychological phenomena, the AB is typically measured by averaging across accuracy scores of multiple participants. In most experiments, all participants show some AB impairment, however, the strength of the AB impairment varies from subject to subject. Hence, the average AB curve 'washes out' these individual differences and only represents the trend underlying the data that is common to the subject population as a whole. A minority of participants, however, seem to not show an AB at all. Such individuals have been termed *non-blinkers* (Feinstein, Stein, Castillo, & Paulus, 2004; Martens, Munneke, et al., 2006) and it is intriguing to investigate why these people seem to be 'immune' to the AB, whereas the rest of the population seem to be *blinkers*. Furthermore, the fact that some people show no AB at all argues against the hypothesis that the AB is a fundamental (or even anatomically) defined limitation of the visual system.

Although the study that recorded EEG profiles for blinkers and non-blinkers during an AB task was already discussed in Section 2.3.2, we will reiterate the main findings. First, Martens, Munneke, et al. (2006) find that non-blinkers, who are behaviourally at ceiling performance throughout the AB, show an earlier P3 component than blinkers. The authors conclude that the non-blinkers' increase in accuracy during the AB might be due to them being faster at consolidating targets into working memory. Second, the frontal selection positivity (FSP) and selection negativity (SN) ERP components are larger for non-blinkers than they are for blinkers. These ERP components have been associated with selective processing of target features (Smid, Jakob, & Heinze, 1999). Consequently, it seems as if non-blinkers are able to employ a more efficient selection mechanism than blinkers, which benefits them when distinguishing between targets and distractors appearing in the RSVP stream. Finally, Martens, Munneke, et al. (2006) find that non-blinkers show less distractor-related EEG activity than blinkers. In line with differences in selection strategies, it seems

as if non-blinkers are more able to suppress distractors while selectively enhancing targets in the RSVP stream.

### Modelling blinkers and non-blinkers using the ST$^2$ model

If we use the ST$^2$ model to hypothesise about blinkers and non-blinkers, it turns out that the model corresponds well with the line of argument from Martens, Munneke, et al. (2006). Although Martens, Munneke, et al. (2006) emphasise how non-blinkers are able to consolidate targets more quickly, the underlying processing differences are hypothesised to be due to a more efficient use of the selection mechanism by non-blinkers.

In the ST$^2$ model, targets are selected and thus distinguished from distractors by means of the task demand mechanism (see Figure 20 and also Section 3.3.2). Since in its normal configuration the ST$^2$ model is designed to produce an AB, it simulates the behavioural accuracy for blinkers. The task demand mechanism operates at the highest layer of stage one (i.e. the task filtered layer) and items have to be extensively processed before task demand can distinguish targets from distractors.

To simulate non-blinkers using the ST$^2$ model, we would have to make a number of theoretically justified changes to the model. In line with Martens, Munneke, et al. (2006), we propose that non-blinkers are better at suppressing individual items appearing in the RSVP stream (i.e. both targets and distractors), but then have a more distinguished task demand mechanism at the task filtered layer. Although greater suppression at early layers means that all items have less activation, the more efficient task demand mechanism could be configured to account for this and allow non-blinkers to consolidate targets in faster and more durable fashion. This would shorten the amount of time it takes to tokenise T1 and increase T2's activation strength, which would reduce the AB.

Greater suppression of individual items in the RSVP stream could be implemented by decreasing the weight between input and masking layers. Activity related to early sensory processing would decrease in this 'non-blinkers simulation' compared to when the model is simulating blinkers. We could then generate virtual ERP components and compare them to the human ERP profiles from Martens, Munneke, et al. (2006). Lower activation at early layers of the ST$^2$ model would be reflected in a reduction of the virtual ssVEP wave (see Section 4.5.1), which would be a qualitative replication of a reduction in average distractor

related EEG activity (see Figure 6 in Martens, Munneke, et al., 2006). Furthermore, the more distinguished task demand mechanism in 'non-blinkers mode' could be implemented by increasing the weight to targets in the task demand layer. This emphasises the difference in activation between targets and distractors at the task demand layer and accelerates the time point of target tokenisation. The effect is an earlier virtual P3 component for non-blinkers compared to blinkers, which matches the P3 effect reported by Martens, Munneke, et al. (2006).

We can make a further prediction for an experimental study comparing the N2pc ERP component for blinkers and non-blinkers. To our knowledge, a study analysing the N2pc component for blinkers and non-blinkers has not been published. However, based on our hypothesised changes to the architecture of the $ST^2$ model so it can simulate non-blinkers, we predict that non-blinkers show an earlier N2pc component than blinkers. Similarly to the P3 component, the increased activation of targets at the task demand layer for non-blinkers compared to blinkers causes the blaster to be triggered at a slightly earlier time point, which decreases the latency of the virtual N2pc component. Hence, this is another example of how the virtual ERP technique can be used to generate testable predictions for future experimental research.

Finally, Martens, Munneke, et al. (2006) present the FSP (frontal selection positivity) and SN (selection negativity) ERP components as correlates of target selection and show that these ERP components are larger for non-blinkers than for blinkers. According to our hypothesis, non-blinkers are able to employ a more distinguished task demand mechanism. Hence, the output from the neurons of the task demand mechanism in the $ST^2$ model could contribute to a virtual FSP/SN component. Future work could explore if the temporal dynamics of this new virtual component replicate the ERP data from Martens, Munneke, et al. (2006) and if this was the case, it would strengthen the hypothesis derived from the $ST^2$ model.

# Appendix A

# Methods

## A.1 Experiment 1

### A.1.1 Participants

Twenty-two under- and postgraduate university students (mean age 22.4; SD 3.2; 10 female; 20 right-handed) provided written consent and received 10 GBP for participation. Two participants were excluded due to an excessive number of EEG artifacts, leaving 18 participants for the behavioural and EEG analysis (mean age 22.2; SD 3.3; 9 female; 19 right-handed). Participants were free from neurological disorders and had normal or corrected-to-normal vision. The study was approved by the local ethics committee.

### A.1.2 Stimuli and apparatus

We presented alphanumeric characters in black on a white background at a distance of 100cm on a 21" CRT computer screen (1024x768 @ 85Hz) using the Psychophysics toolbox (Brainard, 1997) running on Matlab version 6.5 under Microsoft Windows XP. Stimuli were in Arial font and had an average size of 2.1° x 3.4° visual angle. A photodiode verified exact stimulus presentation timing.

### A.1.3 Procedure

Participants viewed four blocks (3 RSVP/1 skeletal, counterbalanced between subjects) of 100 trials. Within each block, there were 96 trials containing a single target and four

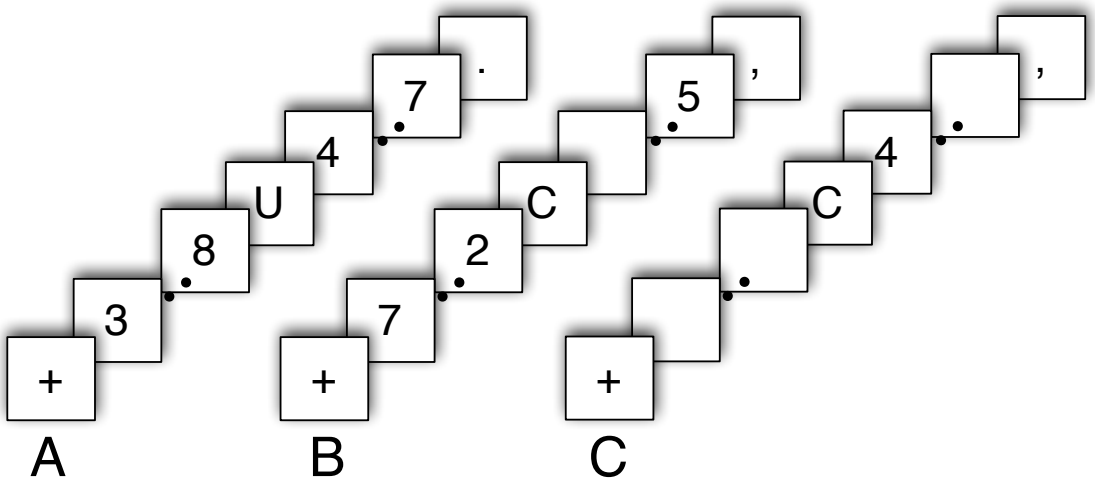


*Figure 58* The experimental paradigm used in Experiment 1. Panel A: An RSVP stream using digits as distractors and a letter as the target. Panel B: A T+1 blank stream where the distractor following the target is omitted. Panel C: A skeletal stream containing only the target letter and the following digit distractor as its mask.

trials consisting only of distractors. Five practice trials preceded the first block in both the RSVP and skeletal conditions, which were not included in the final analysis. The underlying structure and timing of RSVP and skeletal streams were the same; however, whereas in RSVP the target was embedded into a continuous stream of distractors, skeletal streams contained only the target and a following distractor. The target for each trial was chosen at random from a list of 14 capital letters (B, C, D, E, F, G, J, K, L, P, R, T, U, V); distractors could be any digit except 1 or 0. The target item's position in the stream varied between position 10 to 54. The 'distractor only' trials were randomly inserted to make the occurrence of the target less predictable. Trials were randomly ordered and 50% of targets were followed by a blank in both RSVP and skeletal trials to equate patterns within blocks. However, the data from the skeletal unmasked (skeletal stream where the target is not followed by a distractor) condition is not analysed. Figure 58 depicts the three conditions (Panel A: regular RSVP, Panel B: RSVP T+1 blank, Panel C: skeletal presentation) that were analysed from Experiment 1. Figure 58A depicts a single target embedded in a regular RSVP stream. Figure 58B illustrates a T+1 blank stream, where the distractor following the target is removed. Figure 58C shows a skeletal stream consisting solely of the target and the following distractor. Although some studies (Ward et al., 1997) employ patterns

instead of digits to mask the targets, the important difference with respect to RSVP is that all other distractor items are omitted.

A fixation cross presented for 500ms preceded the first item of each stream. Items were presented at the unconventionally fast rate of approx. 20 items per second (item duration 47.1ms; no inter-stimulus interval) to ensure participants' detection accuracy was not at ceiling in this relatively easy single target detection task.

An RSVP stream consisted of 70 items (total stream length 3.3 seconds) to allow a sufficient amount of time between target presentation and the end of the stream. The skeletal condition contained a blank screen for 471ms to 2.5 seconds (depending on the target position), then the target (and its mask in the masked condition) for 47.1ms each, followed by another 706ms to 2.8 seconds of blank screen. The relatively long time period between the presentation of the target and the end of the stream ensured that the subject's behavioural response did not interfere with the EEG signal evoked by the target. Each stream ended with a dot or a comma presented for 47.1ms. Following stream presentation, participants were asked 'Was the final item a comma or a dot?' and in the following screen 'If you saw a letter, type it. If not, press Space.'. Participants entered their responses using a computer keyboard. The dot-comma task was included to ensure that participants maintained their attention on the stream after the target had passed.

### A.1.4 EEG recording

EEG activity was recorded from Ag/Ag-Cl electrodes mounted on an electrode cap (FMS, Munich, Germany) using a high input impedance amplifier (1000MΩ, BrainProducts, Munich, Germany) with a 22-bit analogue-to-digital converter. Electrode impedance was reduced to less than 25kΩ before data acquisition (Ferree, Luu, Russell, & Tucker, 2001). EEG amplifier and electrodes employed actiShield technology (BrainProducts, Munich, Germany) for noise and artifact reduction.

The sampling rate was 2000Hz (digitally reduced to 1000Hz at a later stage) and the data was digitally filtered at low-pass 85Hz and high-pass 0.5Hz during recording. 20 electrodes were placed at the following standard locations according to the international 10/20 system (Jasper, 1958): Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, C7, C8, Pz, P3, P4, P7, P8, Oz, O1, O2, T7 and T8. Electrooculographic (EOG) activity was bipolarly

recorded from below and to the right side of the right eye.

### A.1.5 EEG data analysis

The EEG data was analysed using BrainVision Analyzer (BrainProducts, Munich, Germany), in conjunction with EEGLAB 6.01b (Delorme & Makeig, 2004) and custom MATLAB scripts. The data was referenced to a common average online and re-referenced to linked earlobes offline. Left mastoid acted as ground. Signal deviations in the EOG channel of more than 50$\mu$V within an interval of 100ms were identified as eye blink and movement artifacts. These were removed by rejecting data in the window of 200ms before and after an eye artifact. To verify that these trials were accurately identified by the algorithm, we performed a manual inspection after the algorithm had been applied. ERPs were time locked to the onset of the target and extracted from -200ms to 1200ms with respect to target onset. After segmentation, direct current drift artifacts were removed using a DC detrend procedure employing the average activity of the first and last 100ms of a segment as starting and end point, respectively. Following this, the baseline was corrected to the prestimulus interval (-200ms to time point 0) and segments were averaged to create ERPs.

Unless otherwise stated, ERP component amplitudes were derived from mean amplitude values within a certain window. ERP component latencies were calculated using 50% area latency analysis (Luck & Hillyard, 1990). Amplitude and latency values from subject averages were submitted to Matlab scripts (Trujillo-Ortiz, Hernandez-Walls, & Trujillo-Perez, 2004; Trujillo-Ortiz, Hernandez-Walls, Castro-Perez, & Barba-Rojo, 2006) to perform repeated measures Analysis of Variance (ANOVA). Where appropriate, p-values were adjusted using Greenhouse-Geisser correction. After all statistical analyses, a 25Hz low pass filter was applied to enhance visualisation of ERP components.

### A.1.6 Computational modelling

In order to simulate single target RSVP streams with 50ms presentation rate, the input patterns presented to the $ST^2$ model contained 40 items with the target appearing at position 14 of the stream. Each item was presented for 10 timesteps, which is equivalent to 50ms.

Each item presented to the $ST^2$ model has a certain strength value. Distractors have a constant value of 0.526. To simulate the single target paradigm for Experiment 1, the target strength values iterate from 0.442 to 0.61 in steps of 0.014. This results in the $ST^2$ model simulating 13 trials for the single target paradigm, one simulated trial per target strength.

## A.2 Experiment 2

### A.2.1 Participants

We recruited 20 new under- and postgraduate university students (mean age 23.1, SD 3.2; 10 female; 18 right-handed) who provided written consent and received 10 GBP for participation. Two participants were excluded from the analysis. The first one seemed to be a non-blinker (Martens, Munneke, et al., 2006), as his performance was at ceiling across all three lags. The second participant was excluded due to persistently high oscillations in the alpha band throughout the experiment. Hence, 18 participants remained for behavioural and EEG analysis (mean age 22.5, SD 2.7; 9 female; 18 right-handed). Participants were free from neurological disorders and had normal or corrected-to-normal vision. The study was approved by the local ethics committee.

### A.2.2 Stimuli and apparatus

Stimulus presentation was equal to that in Experiment 1 except for a reduction in average stimulus size (1.03° x 0.69° visual angle) to ensure that the paradigm produced a reliable AB effect.

#### Procedure

Participants viewed four blocks of 100 trials. Before starting the experiment, participants were asked to make 5 eye blinks and 5 horizontal eye movements to record the typical pattern of EOG activity. This was used to configure the algorithm for eye blink artifact rejection. Participants performed 8 practice trials, which were not included in the analysis. As shown in Figure 59, RSVP streams were preceded by a fixation cross in the centre of the

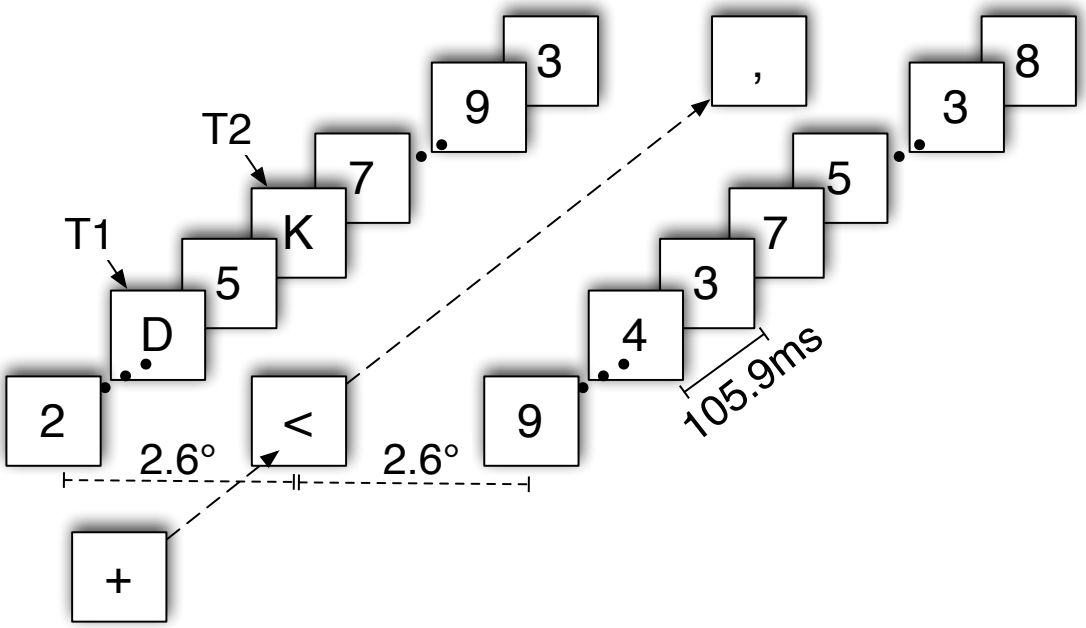


*Figure 59* The two-target bilateral RSVP paradigm used in Experiment 2.

screen. After 400ms, the cross turned into an arrow indicating the side at which the targets would be presented. After 200ms, two streams of digits were simultaneously presented at an equal distance of 2.6° visual angle to the left and right of fixation. The RSVP stream consisted of 35 items presented for 105.9ms each with no inter-stimulus interval. For 84% of trials in a block, the stream on the side indicated by the arrow contained 2 targets (T1 & T2), in 16% of trials both streams were made up of distractor digits only. The 'distractor only' trials were randomly inserted to make the occurrence of targets less predictable. In a trial, T1 and T2 were selected from a list of 18 possible targets (A, B, C, D, E, F, G, H, J, K, L, N, P, R, T, U, V, Y); distractors could be any digit except 1 or 0. T1 appeared between position 5 and 17; T2 followed T1 at position 1 (no intervening distractors - lag 1), position 3 (2 intervening distractors - lag 3) or position 8 (7 intervening distractors - lag 8). The arrow remained in the centre of the screen until the streams were over and then turned into either a dot or a comma.

Before the experiment started, participants were told to keep their eyes fixated on the centre of the screen from presentation of the cross until the dot/comma, as trials with eye movements would be identified in the EOG and excluded from the analysis. Participants

were told to direct their covert attention towards the indicated stream, search for the two target letters and remember whether the last item was a dot or a comma. Participants were informed that streams could contain either two or zero targets. Following stream presentation, participants were presented with the message 'If you saw letters - type them in order, then dot or comma for the final item' and entered their response without time pressure using a computer keyboard. The dot-comma task was included to ensure that participants kept their eyes fixated on the centre of the screen throughout the duration of the RSVP stream.

### A.2.3 EEG recording

For Experiment 2, the sampling rate was 1000Hz and the data was filtered at 80 Hz low-pass and 0.25 Hz high-pass during recording. Horizontal eye movements, recorded from a bipolar EOG channel placed below and to the left of the participant's left eye, indicated that participants had moved their eyes away from fixation and towards one of the RSVP streams.

### A.2.4 EEG data analysis

As with Experiment 1, the EEG data was analysed using BrainVision Analyzer (Brain-Products, Munich, Germany), in conjunction with EEGLAB 6.01b (Delorme & Makeig, 2004) and custom MATLAB scripts. Signal deviations in the EOG channel of more than $50\mu$V within an interval of 100ms were identified as eye blinks and movement artifacts, and a window of 200ms before and after an artifact were marked for rejection. These trials, along with trials violating the artifact rejection procedure described for Experiment 1, were excluded from further analysis. To verify that these artifacts were accurately identified by the algorithm, we performed a manual inspection after the algorithm had been applied.

Unless otherwise stated, ERP component amplitudes were derived from mean amplitude values within a certain window. ERP component latencies were calculated using 50% area latency analysis (Luck & Hillyard, 1990). Amplitude and latency values from subject averages were submitted to Matlab scripts (Trujillo-Ortiz et al., 2004, 2006) to perform

repeated measures Analysis of Variance (ANOVA). Where appropriate, p-values were adjusted using Greenhouse-Geisser correction. After all statistical analyses, a 25Hz low pass filter was applied to enhance visualisation of ERP components.

### A.2.5 Computational modelling

In order to simulate two-target RSVP streams with 100ms presentation rate, the input patterns presented to the $ST^2$ model were comprised of 25 items presented for 20 timesteps (equivalent to 100ms) each. T1 appeared at position 7 in the RSVP stream and T2 followed T1 with 0 to 7 distractors (lags 1 - 8) between the two targets.

Each item presented to the $ST^2$ model has a certain strength value. Distractors have a constant value of 0.526. To simulate the two target paradigm in Experiment 2, the strength values for T1 and T2 iterate from 0.442 to 0.61 in steps of 0.014. Hence, the model simulates 169 target strength combinations, which results in 169 trials in the $ST^2$ model's simulation of the two-target paradigm.

# Bibliography


Arend, I., Johnston, S., & Shapiro, K. (2006). Task-irrelevant visual motion and flicker attenuate the attentional blink. *Psychonomic Bulletin & Review*, *13*(4), 600-607.

Aston-Jones, G., Rajkowski, J., & Cohen, J. (2000). Locus coeruleus and regulation of behavioral flexibility and attention. *Progress in Brain Research*, *126*, 165-182.

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 834–846.

Baars, B. (1988). *A cognitive theory of consciousness*. Cambridge University Press, Cambridge, MA.

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition*, *10*(8), 949–963.

Baillet, S., Mosher, J., & Leahy, R. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, *18*(6), 14-30.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, *103*(2), 449–454.

Barnard, P. (1999). Interacting Cognitive Subsystems: modelling working memory phenomena within a multi-processor architecture. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (p. 298-339). Cambridge University Press, New York.

Barnard, P., & Bowman, H. (2004). Rendering Information Processing Models of Cognition and Affect Computationally Explicit: Distributed Executive Control and the Deployment of Attention. *Cognitive Science Quarterly*, *3*(3), 297-328.

Barnard, P., Scott, S., Taylor, J., May, J., & Knightley, W. (2004). Paying Attention to

Meaning. *Psychological Science*, *15*(3), 179–186.

Battye, G. (2003). *Connectionist Modelling of Attention and Anxiety.* Unpublished doctoral dissertation, The Medical Research Council's Cognition and Brain Sciences unit, Cambridge.

Berger, H. (1929). Ueber das Elektroencephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheit*, *87*, 527-570.

Biederman, I., Rabinowitz, J., Glass, A., & Stacy, W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, *103*, 597-600.

Blankertz, B., Dornhege, G., Krauledat, M., Muller, K.-R., Kunzmann, V., Losch, F., et al. (2006). The Berlin brain-computer interface: EEG-based communication without subject training. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *14*(2), 147-152.

Bledowski, C., Linden, D. E., & Wibral, M. (2007). Combining electrophysiology and functional imaging - different methods for different questions. *Trends in Cognitive Sciences*, *11*(12), 500–502.

Bledowski, C., Prvulovic, D., Goebel, R., Zanella, F., & Linden, D. (2004). Attentional systems in target and distractor processing: a combined ERP and fMRI study. *NeuroImage*, *22*(2), 530–540.

Bledowski, C., Prvulovic, D., Hoechstetter, K., Scherg, M., Wibral, M., Goebel, R., et al. (2004). Localizing P300 Generators in Visual Target and Distractor Processing: A Combined Event-Related Potential and Functional Magnetic Resonance Imaging Study. *Journal of Neuroscience*, *24*(42), 9353-9360.

Botella, J., Barriopedro, M., & Suero, M. (2001). A Model of the Formation of Illusory Conjunctions in the Time Domain. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(6), 1452–1467.

Bowman, H., & Wyble, B. (2007). The Simultaneous Type, Serial Token Model of Temporal Attention and Working Memory. *Psychological Review*, *114*(1), 38-70.

Bowman, H., Wyble, B., Chennu, S., & Craston, P. (2008). A reciprocal relationship between bottom-up trace strength and the attentional blink bottleneck: Relating the LC-NE and $ST^2$ models. *Brain Research*, *1202*, 25–42.

Brainard, D. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433-436.

Breitmeyer, B. G., Ehrenstein, A., Pritchard, K., Hiscock, M., & Crisan, J. (1999). The roles of location specificity and masking mechanisms in the attentional blink. *Perception & Psychophysics*, *61*(5), 179-809.

Breitmeyer, B. G., Ro, T., & Ogmen, H. (2004). A comparison of masking by visual and transcranial magnetic stimulation: implications for the study of conscious and unconscious visual processing. *Consciousness and Cognition*, *13*(4), 829–843.

Broadbent, D. (1958). *Perception and Communication.* London: Pergamon Press.

Broadbent, D., & Broadbent, M. (1987). From Detection to Identification: Response to Multiple Targets in a Rapid Serial Visual Presentation. *Perception & Psychophysics*, *42*, 105–113.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547.

Castiello, U., & Umilta, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 837–848.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *9*(7), 349–354.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory.* Oxford University Press: New York.

Chartier, S., Cousineau, D., & Charbonneau, D. (2004). A Connectionist Model of the Attentional Blink Effect During a Rapid Serial Visual Presentation Task. In L. Earlbaum (Ed.), *Sixth international conference on cognitive modeling* (pp. 64–69).

Chennu, S., Craston, P., Wyble, B., & Bowman, H. (2008). Transient Attentional Enhancement during the Attentional Blink: ERP correlates of the $ST^2$ model. In R. French & E. Thomas (Eds.), *Proceedings of the 10th neural computation and psychology workshop.* London, UK: World Scientific.

Chennu, S., Craston, P., Wyble, B., & Bowman, H. (in revision). Electrophysiological evidence for a reduction in attentional precision during the attentional blink.

Chua, F., Goh, J., & Hon, N. (2001). Nature of Codes Extracted During the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(5), 1229-1242.

Chun, M. (1997a). Temporal binding errors are redistributed by the attentional blink. *Perception & Psychophysics*, *59*(8), 1191-1199.

Chun, M. (1997b). Types and Tokens in Visual Processing: A Double Dissociation Between the Attentional Blink and Repetition Blindness. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(3), 738-755.

Chun, M., & Potter, M. (1995). A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 109–127.

Chun, M., & Wolfe, J. (2001). Visual Attention. In B. Goldstein (Ed.), *Blackwell handbook of perception* (p. 272-310). Oxford, UK: Blackwell Publishers Ltd.

Cohen, J., Dunbar, K., & McClelland, J. (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychological Review*, *97*(3), 332-361.

Cohen, J., Romero, R., Servan-Schreiber, D., & Farah, M. (1994). Mechanisms of spatial attention: The relation of macrostructure to microstructure in parietal neglect. *Journal of Cognitive Neuroscience*, *6*(4), 377-387.

Cooper, A., Humphreys, G., Hulleman, J., Praamstra, P., & Georgeson, M. (2004). Transcranial magnetic stimulation to right parietal cortex modifies the attentional blink. *Experimental Brain Research*, *155*(1), 24–29.

Cooper, R., Fox, J., Farringdon, J., & Shallice, T. (1996). A systematic methodology for cognitive modelling. *Artificial Intelligence*, *85*(1-2), 3–44.

Craston, P., Wyble, B., & Bowman, H. (2006). An EEG study of masking effects in RSVP [Abstract]. *Journal of Vision*, *6*(6), 1016–1016.

Craston, P., Wyble, B., & Bowman, H. (2007). Using virtual ERPs to explore the attentional blink [Abstract]. In *Proceedings of the British Experimental Psychology Society Meeting.* Edinburgh, UK.

Craston, P., Wyble, B., Chennu, S., & Bowman, H. (2009). The attentional blink reveals serial working memory encoding: Evidence from virtual & human event-related potentials. *Journal of Cognitive Neuroscience*, *21*(3), 550-566.

Crick, F., & Koch, C. (1995). Why neuroscience may be able to explain consciousness. *Scientific American*, *273*, 84-85.

David, O., Harrison, L., & Friston, K. J. (2005). Modelling event-related responses in the brain. *NeuroImage*, *25*(3), 756–770.

David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., & Friston, K. J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage*, *30*(4), 1255–1272.

David, O., Kilner, J. M., & Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage*, *31*(4), 1580–1591.

Debener, S., Makeig, S., Delorme, A., & 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), 309–321.

Debener, S., Mullinger, K. J., Niazy, R. K., & Bowtell, R. W. (2008). Properties of the ballistocardiogram artefact as revealed by EEG recordings at 1.5, 3 and 7 T static magnetic field strength. *International Journal of Psychophysiology*, *67*(3), 189-199.

Debener, S., Strobel, A., Sorger, B., Peters, J., Kranczioch, C., Engel, A. K., et al. (2007). Improved quality of auditory event-related potentials recorded simultaneously with 3-T fMRI: Removal of the ballistocardiogram artefact. *NeuroImage*, *34*(2), 587–597.

Debener, S., Ullsperger, M., Siegel, M., & Engel, A. (2006). Single-trial EEG-fMRI reveals the dynamics of cognitive function. *Trends in Cognitive Sciences*, *10*(12), 558–563.

Debener, S., Ullsperger, M., Siegel, M., & Engel, A. K. (2007). Towards single-trial analysis in cognitive brain research. *Trends in Cognitive Sciences*, *11*(12), 502–503.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-Trial Coupling of Concurrent Electroencephalogram and Functional Magnetic Resonance Imaging Identifies the Dynamics of Performance Monitoring. *Journal of Neuroscience*, *25*(50), 11730–11737.

Dehaene, S., Sergent, C., & Changeux, J.-P. (2003). A Neuronal Network Model linking Subjective Reports and Objective Physiological Data during Conscious Perception. *Proceedings National Academie of Sciences*, *100*(14), 8520–8525.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. *PLoS Biology*, *5*(10), e260.

Dell'Acqua, R., Pascali, A., Jolicoeur, P., & Sessa, P. (2003). Four-dot masking produces the attentional blink. *Vision Research*, *43*(18), 1907–1913.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of

single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Descartes, R. (1641). *Meditations on first philosophy.* Paris, France: Michel Soly.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some Theoretical Considerations. *Psychological Review*, *70*(1), 80–90.

Di Russo, F., Teder-Sälejärvi, W., & Hillyard, S. (2003). Steady-state VEP and attentional visual processing. In A. Zani & A. Proverbio (Eds.), *The cognitive electrophysiology of mind and brain* (p. 259-274). San Diego: Academic Press.

Di Lollo, V., Kawahara, J., Ghorashi, S., & Enns, J. (2005). The Attentional Blink: Resource depletion or temporary loss of control? *Psychological Research*, *69*(3), 191-200.

Donchin, E. (1981). Surprise! . . . Surprise? *Psychophysiology*, *18*, 493-513.

Donchin, E., & Coles, M. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357-374.

Donchin, E., Spencer, K., & Wijesinghe, R. (2000). The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *Rehabilitation Engineering, IEEE Transactions on [see also IEEE Trans. on Neural Systems and Rehabilitation]*, *8*(2), 174–179.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, *369*, 313-315.

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225-234.

Engel, A., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences*, *5*(1), 16–25.

Enns, J., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345–352.

Eriksen, C., & Yeh, Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 583-597.

Evans, M., Hastings, N., & Peacock, B. (2000). von Mises Distribution. In *Statistical distributions* (3rd ed., p. 189-191). New York: Wiley.

Feinstein, J., Stein, M., Castillo, G., & Paulus, M. (2004). From sensory processes to

conscious perception. *Consciousness and Cognition*, *13*(2), 323–335.

Fell, J., Fernandez, G., Klaver, P., Elger, C., & Fries, P. (2003). Is synchronized neuronal gamma activity relevant for selective attention? *Brain Research Reviews*, *42*(3), 265–272.

Fell, J., Klaver, P., Elger, C., & Fernández, G. (2002). Suppression of EEG Gamma Activity may cause the Attentional Blink. *Consciousness & Cognition*, *11*(1), 114–122.

Ferree, T. C., Luu, P., Russell, G. S., & Tucker, D. M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology*, *112*(3), 536–544.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1-2), 3–71.

Fragopanagos, N., Kockelkoren, S., & Taylor, J. (2005). A neurodynamic model of the attentional blink. *Cognitive Brain Research*, *24*(3), 568–586.

Frigg, R., & Hartmann, S. (2006). *Models in Science.* Stanford Encyclopedia of Philosophy. Available from `http://plato.stanford.edu/entries/models-science/`

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302.

Garson, J. (2007). *Connectionism.* The Stanford Encyclopedia of Philosophy. Available from `http://plato.stanford.edu/archives/spr2007/entries/connectionism/`

Gazzaniga, M. S. (Ed.). (2004). *The Cognitive Neurosciences III* (3rd ed.). Cambridge, MA, USA: MIT Press.

Giesbrecht, B., Bischof, W., & Kingstone, A. (2003). Visual Masking during the Attentional Blink: Tests of the Object Substitution Hypothesis. *Journal of Experimental Psychology: Human Perception & Performance*, *29*, 238–258.

Giesbrecht, B., Bischof, W., & Kingstone, A. (2004). Seeing the light: Adapting luminance reveals low-level visual processes in the attentional blink. *Brain and Cognition*, *55*(2), 307–309.

Giesbrecht, B., & Di Lollo, V. (1998). Beyond the Attentional Blink: Visual Masking by Object Substitution. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(5), 1454-1466.

Goldstein, J. (1999). Emergence as a Construct: History and Issues. *Emergence: Complexity and Organization*, *1*, 49-72.

Grandison, T., Ghirardelli, T., & Egeth, H. (1997). Beyond similarity: Masking of the target is sufficient to cause the attentional blink. *Perception & Psychophysics*, *59*(2), 266-274.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., et al. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings National Academy of Sciences*, *101*(35), 13050-13055.

Habekost, T., & Bundesen, C. (2003). Patient assessment based on a theory of visual attention (TVA): subtle deficits after a right frontal-subcortical lesion. *Neuropsychologia*, *41*(9), 1171–1188.

Hillstrom, A., Husain, M., Shapiro, K., & Rorden, C. (2004). Spatiotemporal dynamics of Attention in Visual Neglect: A Case Study. *Cortex*, *40*, 433-440.

Hillyard, S., & Münte, T. (1984). Selective attention to color and location: An analysis with event-related brain potentials. *Perception & Psychophysics*, *36*, 185-198.

Hillyard, S., Vogel, E., & Luck, S. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society: Biological Sciences*, *353*, 1257-1270.

Ho, C., Mason, O., & Spence, C. (2007). An investigation into the temporal dimension of the Mozart effect: Evidence from the attentional blink task. *Acta Psychologica*, *125*, 117-128.

Hochstein, S., & Ahissar, M. (2002). View from the Top: Hierarchies and Reverse Hierarchies in the Visual System. *Neuron*, *36*(5), 791–804.

Hodgkin, A., & Huxley, A. (1952). A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. *Journal of Physiology (London)*, *117*, 500-544.

Hommel, B., & Akyürek, E. (2005). Lag-1 Sparing in the Attentional Blink: Benefits and Costs of Integrating two Events into a Single Episode. *Quarterly Journal of Experimental Psychology (A)*, *58*(8), 1415-33.

Hopf, J.-M., Luck, S. J., Girelli, M., Hagner, T., Mangun, G. R., Scheich, H., et al. (2000). Neural Sources of Focused Attention in Visual Search. *Cereb. Cortex*, *10*(12), 1233–1241.

James, W. (1890). *The Principles of Psychology.* New York: Dover.

Jasper, H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography & Clinical Neurophysiology*, *10*, 371-375.

Johnson, R. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, *23*, 367-384.

Johnston, J. C., & McClelland, J. L. (1974). Perception of Letters in Words: Seek Not and Ye Shall Find. *Science*, *184*(4142), 1192–1194.

Jolicoeur, P., Sessa, P., Dell'Acqua, R., & Robitaille, N. (2006). Attentional control and capture in the attentional blink paradigm: Evidence from human electrophysiology. *European Journal of Cognitive Psychology*, *18*(4), 560-578.

Kalidindi, K., & Bowman, H. (2007). Using e-greedy reinforcement learning methods to further understand ventromedial prefrontal patients' deficits on the Iowa Gambling Task. *Neural Networks*, *20*(6), 676-689.

Kanwisher, N. (1987). Repetition Blindness: Type recognition without token individuation. *Cognition*, *27*, 117–143.

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition*, *79*(1-2), 89–113.

Kavanagk, R., Darcey, T., Lehmann, D., & Fender, D. (1978). Evaluation of Methods for Three-Dimensional Localization of Electrical Sources in the Human Brain. *IEEE Transactions on Biomedical Engineering*, *25*(5), 421 - 429.

Kawahara, J., Kumada, T., & Di Lollo, V. (2006). The attentional blink is governed by a temporary loss of control. *Psychonomic Bulletin & Review*, *13*(5), 886-890.

Kenemans, J., Kok, A., & Smulders, F. (1993). Event-related potentials to conjunctions of spatial frequency and orientation as a function of stimulus parameters and response requirements. *Electroencephalography & Clinical Neurophysiology*, *88*, 51-63.

Kessler, K., Schmitz, F., Gross, J., Hommel, B., Shapiro, K., & Schnitzler, A. (2005a). Cortical mechanisms of attention in time: neural correlates of the Lag-1-sparing phenomenon. *European Journal of Neuroscience*, *21*(9), 2563–2574.

Kessler, K., Schmitz, F., Gross, J., Hommel, B., Shapiro, K., & Schnitzler, A. (2005b). Target consolidation under high temporal processing demands as revealed by MEG. *NeuroImage*, *26*(4), 1030-1041.

Kihara, K., Hirose, N., Mima, T., Abe, M., Fukuyama, H., & Osaka, N. (2007). The role of left and right intraparietal sulcus in the attentional blink: a transcranial magnetic

stimulation study. *Experimental Brain Research*, *178*(1), 135–140.

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible'. *Trends in Cognitive Sciences*, *9*(8), 381–388.

Koch, C. (2007). The Neuroscience of Consciousness. In *Fundamental neuroscience, 3-rd edition.* Elsevier.

Koch, C., & Hepp, K. (2006). Quantum mechanics in the brain. *Nature*, *440*(7084), 611–611.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, *38*(3), 557-577.

Kramer, A. F., & Hahn, S. (1995). Splitting the beam:. Distribution of Attention Over Noncontiguous Regions of the Visual Field. *Psychological Science*, *6*(6), 381–386.

Kranczioch, C., Debener, S., & Engel, A. (2003). Event-Related Potential Correlates of the Attentional Blink Phenomenon. *Cognitive Brain Research*, *17*(1), 177–187.

Kranczioch, C., Debener, S., Herrmann, C., & Engel, A. (2006). EEG gamma-band activity in rapid serial visual presentation. *Experimental Brain Research*, *169*(2), 246-254.

Kranczioch, C., Debener, S., Maye, A., & Engel, A. (2007). Temporal dynamics of access to consciousness in the attentional blink. *NeuroImage*, *37*(3), 947-955.

Kranczioch, C., Debener, S., Schwarzbach, J., Goebel, R., & Engel, A. (2005). Neural Correlates of Conscious Perception in the Attentional Blink. *NeuroImage*, *24*(3), 704–714.

Kutas, M., & Hillyard, S. A. (1980). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Brain and Language*, *207*, 203–205.

Kutas, M., & Van Petten, C. (1994). Psycholinguistics Electrified: Event-Related Brain Potential Investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). Academic Press.

Lawrence, D. (1971). Two studies of visual search for word targets with controlled rates of presentation. *Perception & Psychophysics*, *10*(2), 85-89.

Levine, D. S. (2000). *Introduction to Neural and Cognitive Modeling* (2nd ed.; N. Mahwah, Ed.). Lawrence Erlbaum Associates.

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique*. MIT Press, Cambridge, MA.

Luck, S., & Hillyard, S. (1990). Electrophysiological evidence for parallel and serial processing during visual search. *Perception & Psychophysics*, *48*, 603-617.

Luck, S., & Hillyard, S. (1994a). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*, 291-308.

Luck, S., & Hillyard, S. (1994b). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1000–1014.

Luck, S., Vogel, E., & Shapiro, K. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *383*(6601), 616–618.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210.

Makeig, S., Delorme, A., Westerfield, M., Jung, T., Townsend, J., Courchesne, E., et al. (2004). Electroencephalographic Brain Dynamics Following Manually Responded Visual Targets. *PLoS Biology*, *2*(6), e176.

Makeig, S., Westerfield, M., Jung, T., Covington, J., Townsend, J., Sejnowski, T. J., et al. (1999). Functionally Independent Components of the Late Positive Event-Related Potential during Visual Spatial Attention. *Journal of Neuroscience*, *19*(7), 2665–2680.

Marois, R. (2005). Two-timing attention. *Nature Neuroscience*, *8*(10), 1285-1286.

Marois, R., Chun, M., & Gore, J. (2000). Neural Correlates of the Attentional Blink. *Neuron*, *29*, 229–308.

Marois, R., Yi, D., & Chun, M. (2004). The Neural Fate of Consciously Perceived and Missed Events in the Attentional Blink. *Neuron*, *41*, 465–472.

Martens, S., Elmallah, K., London, R., & Johnson, A. (2006). Cuing and stimulus probability effects on the P3 and the AB. *Acta Psychologica*, *123*(3), 204–218.

Martens, S., Munneke, J., Smid, H., & Johnson, A. (2006). Quick Minds Dont Blink: Electrophysiological Correlates of Individual Differences in Attentional Selection. *Journal of Cognitive Neuroscience*, *18*(9), 1423-1438.

Massaro, D., Cohen, M., Campbell, C., & Rodriguez, T. (2001). Bayes factor of model

selection validates FLMP. *Psychonomic Bulletin & Review*, *8*, 1-17.

McArthur, G., Budd, T., & Michie, P. (1999). The Attentional Blink and P300. *Neuroreport*, *10*(17), 3691-3695.

McLaughlin, E., Shore, D., & Klein, R. (2001). The attentional blink is immune to masking-induced data limits. *The Quarterly Journal of Experimental Psychology: Section A*, *54*(1), 169 - 196.

McMains, S., & Somers, D. (2004). Multiple Spotlights of Attentional Selection in Human Visual Cortex. *Neuron*, *42*(4), 677–686.

Moore, C., Egeth, H., Berglan, L., & Luck, S. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review*, *3*(3), 360-365.

Müller, M., & Hillyard, S. (2000). Concurrent recording of steady-state and transient event-related potentials as indices of visual-spatial selective attention. *Clinical Neurophysiology*, *111*(9), 1544–1552.

Müller, M., & Hubner, R. (2002). Can the Spotlight of Attention be Shaped Like A Doughnut? *Psychological Science*, *13*(2), 119-124.

Müller, M., Picton, T., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W., & Hillyard, S. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20-28 Hz range. *Cognitive Brain Research*, *6*(4), 249–261.

Nakatani, C., Ito, J., Nikolaev, R. A., Gong, P., & van Leeuwen, C. (2005). Phase Synchronization Analysis of EEG during Attentional Blink. *Journal of Cognitive Neuroscience*, *17*(12), 1969-79.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, *29*(11), 1631–1647.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. (2005). Decision Making, the P3, and the Locus Coeruleus-Norepinephrine System. *Psychological Bulletin*, *131*(4), 510-532.

Nieuwenhuis, S., Gilzenrat, M., Holmes, B., & Cohen, J. (2005). The role of the locus coeruleus in mediating the attentional blink: A neurocomputational theory. *Journal of Experimental Psychology: General*, *134*(3), 291-307.

Nieuwenhuis, S., van Nieuwpoort, I., Veltman, D., & Drent, M. (2007). Effects of the noradrenergic agonist clonidine on temporal and spatial attention. *Psychopharmacology*, *193*(2), 261-269.

Nieuwenstein, M., & Potter, M. (2006). Temporal Limits of Selection and Memory Encoding. A Comparison of Whole Versus Partial Report in Rapid Serial Visual Presentation. *Psychological Science*, *17*(6), 471–475.

Nieuwenstein, M., Potter, M., & Theeuwes, J. (2009). Unmasking the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 159-69.

Oksama, L., & Hyöna, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, *11*(5), 631–671.

Olivers, C. (2007). The Time Course of Attention: It Is Better Than We Thought. *Current Directions in Psychological Science*, *16*(1), 11-15.

Olivers, C., & Meeter, M. (2008). A Boost and Bounce Theory of Temporal Attention. *Psychological Review*, *115*(4), 836-63.

Olivers, C., & Nieuwenhuis, S. (2005). The Beneficial Effect of Concurrent Task-Irrelevant Mental Activity on Temporal Attention. *Psychological Science*, *16*(4), 265-269.

Olivers, C., & Nieuwenhuis, S. (2006). The Beneficial Effects of Additional Task Load, Positive Affect, and Instruction on the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 364-379.

Olivers, C., van der Stigchel, S., & Hulleman, J. (2007). Spreading the sparing: against a limited-capacity account of the attentional blink. *Psychological Research*, *71*(2), 126–139.

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

Oostenfeld, R., & Delorme, A. (2007). *Localising Independent Components using DIPFIT2.* Available from `http://www.sccn.ucsd.edu/eeglab/dipfittut/dipfit.html`

O'Reilly, R., & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience.* MIT Press.

Pashler, H. (1996). *The Psychology of Attention.* Cambridge, MA: MIT Press.

Pashler, H., & Badgio, P. C. (1985). Visual attention and stimulus identification. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(2), 105–121.

Penny, W. D., Stephan, K. E., Mechelli, A., & Friston, K. J. (2004). Comparing dynamic causal models. *NeuroImage*, *22*(3), 1157–1172.

Pfurtscheller, G., & Lopes da Silva, F. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, *110*(11), 1842-1857.

Picton, T., Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*(2), 127-152.

Pitt, M., Kim, W., & Myung, I. (2003). Flexibility versus generalizability in model selection. *Psychonomic Bulletin & Review*, *10*, 29-44.

Pitt, M., & Myung, I. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, *6*(10), 421–425.

Pitt, M., Myung, I., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, *109*(3), 472–491.

Popper, K. (1959). *The Logic of Scientific Discovery.* London: Hutchinson.

Popple, A., & Levi, D. (2007). Attentional blinks as errors in temporal binding. *Vision Research*, *47*(23), 2973-2981.

Posner, M. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3-25.

Posner, M. (1994). Attention: The Mechanisms of Consciousness. *Proceedings of the National Academy of Sciences*, *91*(16), 7398–7403.

Potter, M. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(5), 509-522.

Potter, M., Chun, M., Banks, B., & Muckenhoupt, M. (1998). Two Attentional Deficits in Serial Target Search: The Visual Attentional Blink and an Amodal Task-Switch Deficit. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(4), 979-992.

Potter, M., Staub, A., & O'Connor, D. (2002). The Time Course of Competition for

Attention: Attention is Initially Labile. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1149-1162.

Pylyshyn, Z., & Storm, R. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179-197.

Rabbitt, P. (1993). Does it all go together when it goes? The nineteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology Section A*, *46*(3), 385–434.

Raymond, J., Shapiro, K., & Arnell, K. (1992). Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860.

Rees, G., Wojciulik, E., Clarke, K., Husain, M., Frith, C., & Driver, J. (2000). Unconscious activation of visual cortex in the damaged right hemisphere of a parietal patient with extinction. *Brain*, *123*(8), 1624–1633.

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review*, *93*(2), 180–206.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*(2), 358–367.

Rolke, B., Bausenhart, K., & Ulrich, R. (2007). Impaired temporal discrimination within the attentional blink. *Perception & Psychophysics*, *69*, 1295-1304.

Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2001). Missed Prime Words within the Attentional Blink evoke an N400 Semantic Priming Effect. *Psychophysiology*, *38*, 165–174.

Ruthruff, E., Johnston, J., Van Selst, M., Whitsell, S., & Remington, R. (2003). Vanishing Dual-Task Interference After Practice: Has the Bottleneck Been Eliminated or Is It Merely Latent? *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 280-289.

Saiki, J. (2003). Feature binding in object-file representations of multiple moving items. *Journal of Vision*, *3*(1), 6–21.

Schalk, G., McFarland, D., Hinterberger, T., Birbaumer, N., & Wolpaw, J. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1034-1043.

Schmolesky, M., Wang, Y., Hanes, D., Thompson, K., Leutgeb, S., Schall, J., et al. (1998). Signal Timing Across the Macaque Visual System. *Journal of Neurophysiology*, *79*(6), 3272–3278.

Seiffert, A., & DiLollo, V. (1997). Low-Level Masking in the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(4), 1061-1073.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *10*(10), 1391-1400.

Sergent, C., & Dehaene, S. (2004). Is Consciousness a Gradual Phenomenon? Evidence for an All-or-Nothing Bifurcation during the Attentional Blink. *Psychological Science*, *15*(11), 720-728.

Sessa, P., Luria, R., Verleger, R., & Dell'Acqua, R. (2006). P3 latency shifts in the attentional blink: Further evidence for second target processing postponement. *Brain Research*, *1137*, 131-139.

Shapiro, K., Caldwell, J., & Sorensen, R. (1997). Personal Names and the Attentional Blink: A Visual "Cocktail Party" Effects. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(2), 504–514.

Shapiro, K., Driver, J., Ward, R., & Sorensen, R. (1997). Priming from the attentional blink: A failure to extract visual tokens but not visual types. *Psychological Science*, *8*, 95-100.

Shapiro, K., Raymond, J., & Arnell, K. (1994). Attention to Visual Pattern Information produces the Attentional Blink in Rapid Serial Visual Presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 357–371.

Shapiro, K., Raymond, J., & Arnell, K. (1997). The attentional blink. *Trends in Cognitive Sciences*, *1*(8), 291–296.

Shapiro, K., Schmitz, F., Martens, S., Hommel, B., & Schnitzler, A. (2006). Resource sharing in the attentional blink. *Neuroreport*, *17*(2), 163-166.

Shore, D., McLaughlin, E., & Klein, R. (2001). Modulation of the attentional blink by differential resource allocation. *Canadian Journal of Experimental Psychology*, *55*(4), 318–324.

Sirevaag, E., Kramer, A., Coles, M., & Donchin, E. (1989). Resource reciprocity: An

event-related brain potentials analysis. *Acta Psychologica*, *70*(1), 77–97.

Slagter, H., Lutz, A., Greischar, L., Francis, A., Nieuwenhuis, S., Davis, J., et al. (2007). Mental Training Affects Distribution of Limited Brain Resources. *PLoS Biology*, *5*(6), e138.

Smid, H., Jakob, A., & Heinze, H.-J. (1999). An event-related brain potential study of visual selective attention to conjunctions of color and shape. *Psychophysiology*, *36*(2), 264–279.

Strobel, A., Debener, S., Sorger, B., Peters, J. C., Kranczioch, C., Hoechstetter, K., et al. (2008). Novelty-and target-processing during an auditory novelty oddball: A simultaneous event-related potential and functional magnetic resonance imaging study. *NeuroImage*, *40*(2), 869-883.

Su, L., Bowman, H., & Barnard, P. (2007). Performance of Reactive Interfaces in Stimulus Rich Environments, Applying Formal Methods and Cognitive Frameworks. In *Electronic Notes in Theoretical Computer Science* (p. 17). The 2nd International Workshop on Formal Methods for Interactive Systems (FMIS 2007).

Svensson, T., Bunney, B., & Aghajanian, G. (1975). Inhibition of both noradrenergic and serotonergic neurons in brain by the $\alpha$-adrenergic agonist clonidine. *Brain Research*, *92*(2), 291–306.

Swartz, B. (1988). Timeline of the history of EEG and associated fields. *Electroencephalography and Clinical Neurophysiology*, *106*, 173-176.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus Specificity of Phase-Locked and Non-Phase-Locked 40Hz Visual Responses in Human. *Journal of Neuroscience*, *16*(13), 4240-4249.

Taylor, J. (2008). *Mind-body problem.* Scholarpedia. Available from `http://www.scholarpedia.org/article/Mind-body_problem`

Taylor, J., & Rogers, M. (2002). A control model of the movement of attention. *Neural Networks*, *15*(3), 309–326.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Trujillo-Ortiz, A., Hernandez-Walls, R., Castro-Perez, A., & Barba-Rojo, K. (2006). *epsGG: Greenhouse-Geisser epsilon. A MATLAB file.* Available from `http://www.mathworks`

.com/matlabcentral/fileexchange/loadFile.do?objectId=12839

Trujillo-Ortiz, A., Hernandez-Walls, R., & Trujillo-Perez, R. (2004). *RMAOV1: One-way repeated measures ANOVA. A MATLAB file.* Available from http://www.mathworks .com/matlabcentral/fileexchange/loadFile.do?objectId=5576

Van Berkum, J., Hagoort, P., & Brown, C. (1999). Semantic Integration in Sentences and Discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657–671.

Verleger, R. (1988). Event-related potentials and cognition: A critique of the context updating hypothesis and an alternative interpretation of P3. *Behavioral Brain Science*, *11*, 343-427.

Visser, T., Zuvic, S., Bishof, W., & Di Lollo, V. (1999). The attentional blink with targets in different spatial locations. *Psychonomic Bulletin & Review*, *6*(3), 432-436.

Vogel, E., & Luck, S. (2002). Delayed Working Memory Consolidation during the Attentional Blink. *Psychonomic Bulletin & Review*, *9*(4), 739–743.

Vogel, E., Luck, S., & Shapiro, K. (1998). Electrophysiological Evidence for a Postperceptual Locus of Suppression During the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1656–1674.

von Helmholtz, H. (1853). Ueber einige Gesetze der Verteilung elektrischer Stroeme in koerperlichen Leitern mit Anwendung auf die tierisch-elektrischen Versuche. *Annalen der Physik und Chemie*, *28*, 173-182.

von Helmholtz, H. (1867). *Treatise on Physiological Optics, J. Southall, transl. (1962)* (3rd ed.). New York: Dover.

Vul, E., Nieuwenstein, M., & Kanwisher, N. (2008). Temporal Selection Is Suppressed, Delayed, and Diffused During the Attentional Blink. *Psychological Science*, *19*(1), 55–61.

Ward, R., Duncan, J., & Shapiro, K. (1996). The Slow Time-Course of Visual Attention. *Cognitive Psychology*, *30*(1), 79–109.

Ward, R., Duncan, J., & Shapiro, K. (1997). Effects of similarity, difficulty, and nontarget presentation on the time course of visual attention. *Perception & Psychophysics*, *59*(4), 593-600.

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual

attention. *Science*, *238*(4828), 778–780.

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science*, *221*(4615), 1080–1082.

Wolfe, J. (1998). Visual Search. In H. Pashler (Ed.), *Attention.* London, UK: University College London Press.

Woodman, G., & Luck, S. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 121 - 138.

Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: Sparing at a costs. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3).

Wyble, B., Craston, P., & Bowman, H. (2006). *Electrophysiological Feedback in Adaptive Human Computer Interfaces* (Technical Report No. 8-06). Computing Laboratory, University of Kent, UK.