Carl von Ossietzky
Universität
Oldenburg

Neurophysiological Basis of Interacting Emotional and Cognitive Processes

Von der Fakultät für Medizin und Gesundheitswissenschaften
der Carl von Ossietzky Universität Oldenburg
zur Erlangung des Grades und Titels eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

angenommene Dissertation

von Frau Katharina Lingelbach
geboren am 29.10.1993 in Backnang

Gutachter und Erstbetreuer: Prof. Dr. Jochem W. Rieger
Zweitbetreuer: Prof. Dr. Christoph S. Herrmann
Weiterer Gutachter: Prof. Dr. Surjo R. Soekadar
Tag der Disputation: 29.10.2025

Acknowledgement

I would like to sincerely thank everyone who has accompanied and supported me during the last five years. I am incredibly fortunate to have so many wonderful people around me.

First and foremost, I would like to express my deepest gratitude to my primary supervisor, Prof. Dr. Jochem W. Rieger, who has continuously supported and mentored me, both during my dissertation and beyond, in my professional career. Our numerous discussions, his expertise, critical feedback, and encouragement have been fundamental to my growth.

I would also like to thank my second supervisor, Prof. Dr. Christoph S. Herrmann, for his supervision and valuable feedback, particularly in the final months. I am equally and sincerely grateful to Prof. Dr. med. Surjo R. Soekadar for his insightful feedback and expertise, and for serving as an external evaluator and additional member of my thesis committee.

My gratitude also goes to Dr. Tina Schmitt, Dr. Andreas Spiegler, and the members of the Neuroimaging Unit, who supported me with their time and resources during my data collection at the Carl von Ossietzky University Oldenburg. Further thanks go to the PhD community in Oldenburg, who welcomed me warmly and included me in socials and research discussions – particularly to Maren, with whom I found a kindred spirit and dear friend.

Last but not least, a big thank you goes to my team, Applied Neurocognitive Systems, that is, to Anna, Ravi, Nektaria, Michael, and, particularly, my team leader, Dr. Mathias Vukelić. Thank you for being the best team one can wish for, for the lunch discussions on psychology, neuroscience, and philosophy, our wonderful social events, and for being more than colleagues to me. This doctoral thesis, with its deep dive into foundational neuroscience, would not have been possible without your support, understanding, and generosity.

Finally, I most profoundly thank my parents and parents-in-law, particularly my mom, Karin, who travelled with me along stony paths beyond the academic pursuit in the last year, and my mother-in-law, Jutta. Further, I want to thank my brother, Christian, and his lovely family, my sisters-in-law, Lara and Sophie, and my best friends, Nora and Elisabeth. You all are my Bridge 4 and have accompanied me through the ups and downs of the past years with love, optimism, care packages, and laughter. I express my deepest thanks to the most important member of Bridge 4 – my husband, Yannick – for his unwavering support, love, and endless patience – especially when listening to my ongoing ramblings about face processing, brain regions (with cryptic abbreviations), and, of course, the interaction of emotion and cognition.

Research is formalised curiosity.
It is poking and prying with a
purpose.

– ZORA NEALE HURSTON

Summary

The ability to perceive and interpret one's own emotions, as well as those of others, is fundamental to social interaction and adaptive decision-making. Yet emotional information can be a double-edged sword. When aligned with our goals, it enhances cognition and behaviour. However, if the emotional information is irrelevant or conflicting, it can have detrimental effects.

While ancient Greek philosophers already recognised the intricate interplay between emotion and cognition, neuroscientific research has primarily investigated these two phenomena separately. Despite recent advances, our understanding of how emotional and cognitive processes interact and integrate at the neural level remains limited. This is particularly true regarding the influence of positive emotional stimuli and the interaction between emotional valence and varying cognitive load in dual-task paradigms.

This thesis extends the field of research by investigating the interaction and integration of emotional and cognitive processes within dual-task paradigms in which emotional stimuli become temporarily task-relevant (Research Question I). Furthermore, it examines the circumstances in which the valence of emotional information and the current level of cognitive load together influence information processing and integration (Research Question II). The research questions were investigated in two research projects comprising three empirical studies. Project I focused on naturalistic emotional facial expressions and working memory (Study I). Project II concentrated on naturalistic emotional speech (Study II) during simulated driving (Study III).

Study II contributes a novel, validated, and publicly accessible database of naturalistic negative, positive, and neutral speech. As the first German database of contextually coherent speech sequences for emotion induction, it addresses a critical gap in experimental resources for research on emotion processing. A behavioural validation study (Study II) quantified quality criteria of emotion differentiation, salience and strength, as well as inter-individual generalisability. Each speech sequence is accompanied by a wide range of published annotations, including continuous labels capturing the temporal dynamics and variability of perceived valence and arousal. Additionally, discrete emotions, emotional intensity, dominance, and familiarity were assessed after the presentation of each sequence.

The two other empirical studies (Study I in Research Project I and Study III in Research Project II) utilised magnetoencephalography (MEG) and eye-tracking within dual-task paradigms. The studies incorporated both emotional and cognitive task components in the dual tasks.

The first MEG–eye-tracking study (Study I) examined how facial expressions and visuo-spatial working memory load modulate attention allocation and information processing within a sequential dual task. During the emotional task, participants had to discriminate between negative, positive, and neutral facial expressions while retaining information about the cognitive task in their working memory. A working memory encoding and retrieval phase of the cognitive n-back task followed, in which the previously presented emotional face reappeared as a distractor. The second MEG–eye-tracking study (Study III) investigated gaze and oscillatory

signatures of emotional speech processing and visuo-spatial workload during the simultaneous dual task of driving while listening to speech. The emotional speech stimuli were taken from the newly developed GAUDIE database (Study II).

In Study I, permutation-based spatiotemporal clustering in source space revealed increased amplitudes of event-related magnetic fields during the discrimination of negative, as compared to neutral and positive, facial expressions. Amplitudes were not modulated when these faces acted as distractors. The modulations were localised in occipital regions involved in processing faces, as well as fronto-temporal regions responsible not only for recognising emotional expressions, but also for generating and regulating emotional responses. At the end of the discrimination phase, at approximately 800 ms after face onset, late evoked responses increased in regions associated with emotion regulation. The observation implies that these regulatory regions are involved in integrating emotion and cognition, possibly by disengaging from emotional effects and reallocating attentional resources to the subsequent cognitive task. In Study III, source-space spatial clustering identified antagonistic signatures of top-down guided and bottom-up stimulus-driven processing. Results indicate that beta- and gamma-band oscillations support predictive listening during emotional speech processing and sensorimotor processing during visuo-spatial workload. Additionally, a multivariate pattern analysis using common spatial patterns demonstrated that fronto-temporal gamma-band activity is influenced by both emotional speech valence and driving-related workload. In emotion-cognition dual tasks, fronto-temporal gamma oscillations appear to coordinate emotional and cognitive processes by signalling a need for regulation and inhibiting emotional responses based on the combined demand for resources imposed by emotional speech valence and cognitive load.

Taken together, converging evidence across studies suggests that fronto-temporal brain regions integrate emotional and cognitive processes by regulating emotional responses as a function of available cognitive resources (Research Question I). In line with existing theories, the regulatory processes seem to involve probing and simulating emotional experiences, labelling them, distancing oneself from them through self-other distinction, as well as evaluating their value in the current context. Notably, the current work provides novel evidence for a central role of the anterior insula and orbitofrontal cortex in integrating and regulating emotional information during dual tasks. These regions are implicated in the value evaluation of emotional stimuli in relation to the context, demands, and goals of other cognitive processes.

Regarding the joint influence of emotional stimulus valence and cognitive load (Research Question II), both studies suggest that emotion-specific processing primarily occurs under low cognitive load during perceptual tasks. Reduced pupil dilation in response to positive facial expressions and positive subjective experiences during emotional speech implied facilitated discriminative processing and deeper emotional appraisal of positive stimuli under low, but not high, cognitive load. However, the emotion-specific processing under low cognitive load did not interfere with cognitive task performance. By contrast, negative stimuli under high, but not low, cognitive load elicited interference in tasks involving executive functions and working memory. These interferences are likely explained by limited cognitive resources and were reflected in reduced target-directed gaze behaviour (Study I) and diminished gamma-based regulatory processes (Study III). The results offer a novel insight, suggesting that the nature of the task (i.e., perceptual vs. executive task elements) also mediates

the co-modulation of emotional valence and cognitive load. By identifying task type as a factor influencing co-modulatory effects, the findings challenge previous conclusions from electrophysiological and neuroimaging research. These earlier studies had attributed the effects to the source of cognitive load (perceptual vs. working memory load).

Enhanced stimulus-driven and compensatory processing modes likely prevented severe task interference in the cognitive task during high load in both studies. These modes showed reduced temporo-occipital alpha-band power during working memory encoding (Study I), as well as enhanced spatial and sensorimotor processing through decreased beta and increased gamma oscillations during driving (Study III).

In summary, the findings substantially advance the understanding of how emotional valence and cognitive load co-modulate information processing and regulatory control by: (1) employing dual-task paradigms that integrate emotional and cognitive task components, (2) incorporating emotionally negative, neutral, and positive naturalistic stimuli across varying levels of cognitive load, and (3) disentangling the cognitive load manipulation from perceptual task elements.

Further research is necessary to understand how dual-task parameters, such as varying temporal proximity or task instruction, influence emotion-cognition interactions and result in task interference. Moreover, future studies should replicate the current findings in even more naturalistic settings by incorporating dynamic, multisensory emotional stimuli and real-life social interactions.

Insights into the neurophysiological basis of emotion-cognition interaction hold strong translational potential for advancing adaptive assistive systems in safety-critical and performance-driven environments, as well as guiding clinical interventions.

Zusammenfassung

Die Fähigkeit, die eigenen Emotionen und die Emotionen anderer wahrzunehmen, bildet die Grundlage für soziale Interaktionen und adaptive Entscheidungsprozesse. Emotionale Informationen erweisen sich dabei als zweischneidiges Schwert: Sind sie zielrelevant, unterstützen sie adaptive Kognition und Handlungssteuerung; fehlt die Relevanz oder entstehen Zielkonflikte, können sie die kognitive Verarbeitung beeinträchtigen.

Bereits die griechischen Philosophen der Antike erkannten die komplexe Wechselwirkung zwischen Emotion und Kognition. Die neurowissenschaftliche Forschung betrachtete diese beiden Phänomene dagegen lange Zeit weitgehend getrennt. Trotz jüngerer Erkenntnisse über die Zusammenhänge bleibt unser Verständnis sowohl der Interaktion als auch der neuronalen Integration emotionaler und kognitiver Prozesse bis heute begrenzt. Dies gilt vor allem für die Effekte positiver emotionaler Stimuli sowie für das Zusammenspiel der Valenz emotionaler Reize mit variierender kognitiver Beanspruchung in Doppelaufgaben.

Die Dissertation erweitert die aktuellen Forschungserkenntnisse, indem sie die Interaktion und neuronale Integration emotionaler und kognitiver Prozesse in Doppelaufgaben untersucht, in denen emotionale Reize zeitweise aufgabenrelevant sind (Forschungsfrage I). Zudem erforscht sie, unter welchen Bedingungen die Valenz emotionaler Informationen in Kombination mit der aktuellen kognitiven Belastung die Informationsverarbeitung und -integration beeinflusst (Forschungsfrage II). Die beiden Forschungsfragen wurden in zwei Forschungsprojekten mit insgesamt drei empirischen Studien untersucht. Projekt I fokussierte sich auf naturalistische emotionale Gesichtsausdrücke und das Arbeitsgedächtnis (Studie I). Projekt II konzentrierte sich auf naturalistische emotionale Sprache (Studie II) während des simulierten Fahrens (Studie III).

Die Entwicklung und Validierung einer öffentlich zugänglichen Datenbank naturalistischer Sprachstimuli mit positiver, neutraler und negativer Valenz in Studie II leistete einen entscheidenden Beitrag: Sie stellt die erste deutschsprachige, inhaltlich kohärente und validierte Sammlung emotionaler Sprachsequenzen dar und schafft damit eine neue experimentelle Ressource für die Emotionsforschung.

Eine Verhaltensstudie (Studie II) diente der Validierung in Bezug auf Salienz, emotionale Intensität und interindividuelle Generalisierbarkeit der Stimuli. Zu jeder Sequenz sind detaillierte Annotationen veröffentlicht. Die Annotationen umfassen kontinuierliche Bewertungen des zeitlichen Verlaufs und der Variabilität der wahrgenommenen Valenz und Erregung. Ergänzend wurden nach der Sequenzpräsentation diskrete Emotionen, die emotionale Intensität, Dominanz und Vertrautheit erfasst.

In den beiden weiteren empirischen Studien (Studie I im Forschungsprojekt I und Studie III im Forschungsprojekt II) kamen Magnetenzephalographie (MEG) und Eye-Tracking in Doppelaufgaben-Paradigmen zum Einsatz. Beide Studien kombinierten emotionale und kognitive Aufgabenelemente in den Doppelaufgaben.

Zur Beantwortung der Forschungsfragen untersuchte die erste MEG–Eye-Tracking-Studie (Studie I), wie verschiedene Mimiken bei variierender visuo-räumlicher Arbeitsgedächtnisbelastung die Aufmerksamkeit und Informationsverarbeitung in einer sequenziellen Doppelaufgabe beeinflussen. Während der emotionalen

Aufgabe mussten die Probanden zwischen emotionalen und neutralen Gesichtsausdrücken unterscheiden, während sie zeitgleich Informationen zur kognitiven Aufgabe im Arbeitsgedächtnis präsent hielten. Anschließend folgte die Phase der Kodierung neuer sowie die Abfrage gespeicherter Arbeitsgedächtnisinhalte im Rahmen einer N-Back–Aufgabe, bei der das zuvor gezeigte emotionale Gesicht abermals als Ablenkungsreiz präsentiert wurde. Die zweite MEG–Eye-Tracking-Studie (Studie III) untersuchte Blickverhalten und oszillatorische Hirnaktivität während der Verarbeitung emotionaler Sprache und visuell-räumlicher kognitiver Belastung. Dies wurde im Kontext einer realnahen Doppelaufgabe erforscht: simuliertes Autofahren bei gleichzeitigem Hören emotionaler Sprachsequenzen. Die emotionalen Sprachstimuli stammten aus der neu entwickelten GAUDIE-Datenbank (Studie II).

In Studie I zeigten permutation-basierte räumlich-zeitliche Clusteranalysen im Quellenraum erhöhte Amplituden ereignisbezogener Magnetfelder während der Unterscheidung negativer im Vergleich zu positiven und neutralen Gesichtsausdrücken. Dieser Effekt blieb aus, wenn die Gesichtsausdrücke als Ablenkungsreize präsentiert wurden. Die Modulationen waren sowohl in okzipitalen Arealen lokalisiert, die an der Gesichtsverarbeitung beteiligt sind, als auch in fronto-temporalen Regionen, die mit der Erkennung emotionaler Ausdrücke sowie mit der Generierung und Regulation emotionaler Reaktionen assoziiert sind. Am Ende der Diskriminationsphase, etwa 800 ms nach dem Erscheinen des Gesichtsreiz, deuteten späte erhöhte Amplituden ereignisbezogener Magnetfelder in regulatorischer Regionen auf eine Integration emotionaler und kognitiver Prozesse hin. Dies ermöglichte wahrscheinlich sowohl die Loslösung von emotionalen Effekten als auch die gezielte Bereitstellung von Aufmerksamkeitsressourcen für die nachfolgende kognitive Aufgabe. In Studie III wurden im Quellenraum räumliche Cluster identifiziert, die antagonistisch wirkende Signaturen top-down-gesteuerter und bottom-up-stimulusinduzierter Verarbeitung abbildeten. Die Ergebnisse deuten darauf hin, dass das Zusammenwirken von Beta- und Gamma-Band-Oszillationen sowohl das antizipative Zuhören während der emotionalen Sprachverarbeitung als auch die sensomotorische Integration bei visuell-räumlicher kognitiver Belastung fördert. Eine multivariate Analyse unter Verwendung gemeinsamer räumlicher Muster zeigte darüber hinaus, dass die Gamma-Band-Aktivität in fronto-temporalen Regionen durch beide Faktoren – die Valenz der Sprachreize und die fahrbezogene kognitive Belastung – gemeinsam moduliert wird. Die Ergebnisse legen nahe, dass fronto-temporale Gamma-Oszillationen als Koordinatoren zwischen emotionalen und kognitiven Prozessen fungieren: Sie kommunizieren Regulierungsbedarf und hemmen emotionale Reaktionen in Situationen, in denen emotionale Sprachvalenz und kognitive Belastung gemeinsam erhöhte Aufmerksamkeitsressourcen erfordern.

Zusammenfassend belegen die konvergierenden Studienbefunde, dass fronto-temporale Hirnregionen zur Integration emotionaler und kognitiver Prozesse beitragen, indem sie emotionale Reaktionen in Abhängigkeit von den verfügbaren kognitiven Ressourcen regulieren (Forschungsfrage I). Die Ergebnisse dieser Arbeit sprechen dafür, dass die neuronale Regulation – im Einklang mit etablierten Theorien – Prozesse der Simulation und Benennung emotionaler Erfahrungen, der Distanzierung durch die Unterscheidung zwischen Selbst und Anderen sowie der Relevanzbewertung emotionaler Reize im aktuellen Kontext umfasst. Die vorliegende Arbeit liefert erstmals empirische Evidenz für die zentrale Rolle der anterioren Insula und des orbitofrontalen Kortex bei der Integration und Regulation emotionaler Informationen während der zeitgleichen Ausführung emotionaler

und kognitiver Aufgaben. Beide Regionen bewerten emotionale Reize im Hinblick auf den situativen Kontext, die aktuellen Anforderungen und die Zielsetzung anderer kognitiver Prozesse.

Bezüglich des gemeinsamen Einflusses von emotionaler Stimulusvalenz und kognitiver Belastung (Forschungsfrage II) deuten die Ergebnisse beider Studien darauf hin, dass emotionsspezifische Verarbeitungsprozesse insbesondere unter niedriger kognitiver Belastung in Wahrnehmungsaufgaben stattfinden. Eine verringerte Pupillenerweiterung bei positiven Gesichtsausdrücken und positivere subjektive Bewertungen positiver Sprachreize sprechen für eine erleichterte Verarbeitung während der Emotionsdiskrimination sowie für eine tiefere Reizbewertung positiver Stimuli bei geringer, nicht jedoch bei hoher, kognitiver Belastung. Die emotionsspezifische Verarbeitung unter niedriger kognitiver Belastung führte gleichwohl nicht zu Interferenzen und einer Leistungsbeeinträchtigung in der kognitiven Aufgabe. Unter hoher, nicht jedoch geringer, kognitiver Belastung in Aufgaben, die Exekutivfunktionen und das Arbeitsgedächtnis beanspruchen, führten negative Stimuli zu Interferenzen. Die Interferenzen sind mit hoher Wahrscheinlichkeit auf begrenzte kognitive Ressourcen zurückzuführen und manifestierten sich in einem weniger zielgerichteten Blickverhalten (Studie I) sowie in einer reduzierten gamma-basierten Regulationsaktivität (Studie III). Die Befunde zeigen, dass die Art der Aufgabe (Wahrnehmung versus exekutive Komponenten) ebenfalls die Wechselwirkung zwischen emotionaler Valenz und kognitiver Belastung beeinflusst. Die Schlussfolgerung, dass der Aufgabentyp diese Wechselwirkung moduliert, stellt frühere Erklärungsansätze aus der elektrophysiologischen und bildgebenden Forschung infrage. Diese führten Unterschiede bislang auf die Art der kognitiven Beanspruchung zurück (perzeptuelle Last vs. Arbeitsgedächtnislast).

Verstärkte stimulusgesteuerte und kompensatorische Verarbeitungsstrategien verhinderten vermutlich eine stärkere Aufgabeninterferenz bei hoher kognitiver Belastung in beiden Studien. Die Kompensation äußerte sich in einer verminderten temporo-okzipitalen Alpha-Band-Aktivität während der Arbeitsgedächtniskodierung (Studie I) sowie in einer gesteigerten räumlichen und sensomotorischen Verarbeitung, die sich in verringerten Beta- und erhöhten Gamma-Oszillationen während der Fahraufgabe (Studie III) widerspiegelten.

Insgesamt erweitern die Ergebnisse das Verständnis darüber, wie emotionale Valenz und kognitive Belastung gemeinsam die Informationsverarbeitung und die regulatorische Kontrolle beeinflussen. Dies wurde ermöglicht durch (1) die Verwendung von Doppelaufgaben, die emotionale und kognitive Aufgabenelemente integrieren, (2) die Einbeziehung negativer, neutraler und positiver naturalistischer Reize bei variierenden kognitive Belastungsstufen sowie (3) die unabhängige Manipulation der kognitiven Belastung getrennt von Wahrnehmungsaufgaben.

Weitere Forschung sollte untersuchen, inwiefern spezifische Parameter von Doppelaufgaben – etwa die zeitliche Nähe zwischen den Teilaufgaben oder die Art der Aufgabeninstruktion – die Interaktion zwischen Emotion und Kognition sowie das Ausmaß der Aufgabeninterferenzen beeinflussen. Darüber hinaus sollten künftige Studien die vorliegenden Befunde in noch realitätsnäheren Kontexten replizieren, beispielsweise durch die Einbindung dynamischer, multisensorischer emotionaler Reize und realer sozialer Interaktionen.

Ein vertieftes Verständnis der neurophysiologischen Grundlagen der Emotions-Kognitions-Interaktion eröffnet neue Perspektiven für die Anwendung und Weiterentwicklung adaptiver Assistenzsysteme in sicherheitskritischen und leistungsorientierten Umgebungen sowie für die Gestaltung klinischer Interventionen.

Contents

Summary v

Zusammenfassung ix

List of Abbreviations xvii

List of Figures xxv

List of Tables xxvii

1 General Introduction 1

1.1 The Theoretical Basis of Interacting Emotional and Cognitive Processes . . . 2
1.2 The Neurophysiological Basis of Interacting Emotional and Cognitive Processes . . . 9
1.3 Magnetoencephalography . . . 12
1.4 Studying the Emotion-Cognition Interaction in Naturalistic Contexts . . . 15
1.5 Scope and Objectives of the Thesis . . . 15

2 Study I: Neurophysiological Basis of Emotional Face Perception and Working Memory Load 19

2.1 Abstract . . . 21
2.2 Introduction . . . 22
2.3 Methods and Materials . . . 29
2.4 Results . . . 38
2.5 Discussion . . . 46
2.6 Conclusion . . . 52
2.7 Funding . . . 53
2.8 Declaration of Competing Interests . . . 53
2.9 Data and Code Availability . . . 54
2.10 Acknowledgements . . . 54

3 Study II: Development, Validation, and Exploration of a Naturalistic AUDItory Emotional Database 67

3.1 Abstract . . . 69
3.2 Introduction . . . 70
3.3 Methods . . . 72
3.4 Results . . . 77
3.5 Discussion . . . 86

3.6 Conclusion . . . 87
3.7 Funding . . . 87
3.8 Declaration of Competing Interests . . . 87
3.9 Data and Code Availability . . . 87
3.10 Acknowledgements . . . 88

4 Study III: Oscillatory Signatures of Emotional Speech Processing and Visuo-Spatial Cognition 93
4.1 Abstract . . . 95
4.2 Introduction . . . 96
4.3 Methods . . . 100
4.4 Results . . . 111
4.5 Discussion . . . 121
4.6 Conclusion . . . 128
4.7 Funding . . . 129
4.8 Declaration of Competing Interests . . . 129
4.9 Data and Code Availability, Ethics Approval, and Consent of Participation and Publication . . . 129
4.10 Acknowledgements . . . 129

5 General Discussion 141
5.1 The Regulatory Fronto-Temporal Axis in Emotion-Cognition Dual Tasks . . . 142
5.2 Joint Influences of Emotional Valence and Cognitive Load . . . 148
5.3 Limitations and Ongoing Challenges . . . 151
5.4 Future Directions of the Research Field . . . 153
5.5 Practical Implications and Translational Potential . . . 154
5.6 Concluding Remarks . . . 156

General References 159

6 Supplementary Material for Study I 175
6.1 Supplementary Figure of the Experimental Procedure . . . 176
6.2 Summary of the Behavioural Linear Mixed-Effects Models . . . 176
6.3 Spatiotemporal *F*-Statistic of the Permutation-Based Clustering . . . 178
6.4 Exploratory Correlation Analyses of Event-Related Magnetic Fields Clusters . . . 178
6.5 Exploratory Correlation Analyses of Evoked Oscillatory Power and Memory Performance . . . 182
6.6 Summary of the Gaze-Related Linear Mixed-Effects Models . . . 184
6.7 Effects of Dual-Task Phase on Gaze Behaviour . . . 186

7 Supplementary Material for Study II 191
7.1 Detailed Information on Audio Sequences . . . 192
7.2 Sample Characteristics: Personality and Well-Being . . . 197

7.3 Study Instruction . . . 200
7.4 Rating Scales Used in the Validation Study . . . 202
7.5 Supplementary Tables of the SAM Ratings . . . 203
7.6 Probabilities of Cluster Affiliation . . . 208
7.7 Familiarity Ratings . . . 211
7.8 Discrete Emotion Ratings . . . 213
7.9 Alternative Analysis of the Cohen's D as Measure of Discriminatory Power . . . 226

8 Supplementary Material for Study III **229**
8.1 Stimulus Set From the German AUDItory Emotional Database . . . 230
8.2 Analysis of Conversation Content Recall and Misses . . . 230
8.3 Centre Frequencies of the Frequency Bands . . . 231
8.4 Common Spatial Patterns for Multiclass Decoding . . . 231
8.5 Calculation of Gaze Dispersion, Pupil Diameter in Millimetres, and Pupillary Activity Index . . . 231
8.6 Detailed Summary of the Linear Mixed-Effects Models . . . 233
8.7 Aperiodic Component Localised in Source Space With DICS . . . 237
8.8 Modulations in the Non-Corrected Band Source Power Localised With DICS . . . 237
8.9 Supplementary Figures . . . 239

Publication List **253**

Declaration of Academic Integrity **255**

Declaration of Generative AI in Scientific Writing **257**

Curriculum Vitae **259**

List of Abbreviations

Anatomical Brain Regions

ACC	Anterior cingulate cortex
dlPFC	Dorsolateral prefrontal cortex
FFA	Fusiform face area
FO	Frontal operculum
IFG	Inferior frontal gyrus
IPL	Inferior parietal lobule
mPFC	Medial prefrontal cortex
OFA	Occipital face area
OFC	Orbitofrontal cortex
pPFC	Posterior prefrontal cortex
PFC	Prefrontal cortex
SMA	Supplementary motor area
STG	Superior temporal gyrus
STS	Superior temporal sulcus
TP	Temporal pole
TPJ	Temporo-parietal junction
vlPFC	Ventrolateral prefrontal cortex
vmPFC	Ventromedial prefrontal cortex

Neuroanatomical Coordinates and Atlases

AAL Automated Anatomical Labelling (atlas)

BA Brodmann area (atlas)

Hemi Hemisphere

lh Left hemisphere

LPA Left pre-auricular point

mid Middle

MNI Montreal Neurological Institute (coordinate space)

rh Right hemisphere

RPA Right pre-auricular point

sup Superior

Measurement Modalities and Associated Hardware

ECoG Electrocorticography

EEG Electroencephalography

fMRI Functional magnetic resonance imaging

fNIRS Functional near-infrared spectroscopy

HPI Head position indicator (coils)

iEEG Intracranial electroencephalography

MEG Magnetoencephalography

MRI Magnetic resonance imaging

SQUIDs Superconducting quantum interference devices

Experimental Questionnaires

BFI-K	German Big Five Inventory - Short scale
GEW	Geneva Emotion Wheel
NASA-TLX	NASA Task Load Index
SAM	Self-Assessment Manikin
STAI	State-Trait Anxiety Inventory
SWLS	Satisfaction With Life Scale
WHO-5	WHO-5 Well-Being Index

Experimental Parameters

ED	Emotion discrimination
HV	High valence (positive emotion)
HW - Study I	High working memory load
HW - Study II	High visuo-spatial workload
LV	Low valence (negative emotion)
LW - Study I	Low working memory load
LW - Study II	Low visuo-spatial workload
NV	Neutral valence (neutral emotion)
WM	Working memory

Neurocognitive Components and Constructs

ERF	Event-related magnetic field
ERP	Event-related potential
IPA	Index of pupillary activity, also referred to as the index of cognitive activity
LPP	Late positive potential ($\sim$ 300 – 1000 ms) sensitive to salience
N/M170	Occipito-temporal negative/magnetic deflection ($\sim$ 170 ms) sensitive to faces and expertise
P1/P100	Early occipital positivity ($\sim$ 80 – 130 ms) sensitive to early visual processing
P3	Positive deflection ($\sim$ 250 – 500 ms with peak at $\sim$ 300 ms) sensitive to context updates
P3b	Parietal subcomponent of the P3 sensitive to task-relevant targets
ToM	Theory of mind

Stimulus Databases

eNTERFACE'05	Audio-visual emotion database (English)
GAUDIE	German AUDItory Emotional Database (German)
HEU	Natural video emotion database (facial expressions, body posture, speech)
IAPS	International Affective Picture System (images)
IEMOCAP	Interactive Emotional Dyadic Motion Capture Database (English)
LANG	Leipzig Affective Norms for German (words)
RAMAS	Russian Acted Multimodal Affective Set
RAVDESS	Ryerson Audio-Visual Database of Emotional Speech and Song (English)
VAM	Vera am Mittag Database (German)

Statistical Methods and Descriptors

α	Threshold for statistical significance
ANOVA	Analysis of variance
AUC	Area under the curve
CI	Confidence interval
Cronbach's α	Measure of internal consistency
FDR	False discovery rate
JAD	Joint approximate diagonalisation
LDA	Linear discriminant analysis
LMM	Linear mixed-effects model
M	Mean
max	Maximum
MCS	Monte Carlo simulation
min	Minimum
ML	Machine learning
MVPA	Multivariate pattern analysis
r_p	Pearson correlation coefficient
r_s	Spearman rank correlation coefficient
RMS	Root mean square
RMSSD	Root mean square of successive differences
rmANOVA	Repeated-measures analysis of variance
SD	Standard deviation
SNR	Signal-to-noise ratio

Signal Processing and Imaging Parameters

BEM	Boundary element model (source reconstruction)
BOLD	Blood-oxygen-level-dependent (functional MRI)
CSP	Common spatial pattern
CSD	Cross-spectral density
DICS	Dynamic imaging of coherent sources
DPSS	Discrete prolate spheroidal sequence (multitaper for spectral analysis)
DWT	Discrete wavelet transform
FASTER	Fully Automated Statistical Thresholding for EEG Artefact Rejection
FEM	Finite element model (source reconstruction)
FIR	Finite impulse response (filter)
FOOOF	Fitting oscillations and one over f (M/EEG spectral analysis)
GRAPPA	GeneRalized Autocalibrating Partial Parallel Acquisition (MRI)
ICA	Independent component analysis
ICP	Iterative closest point algorithm (co-registration)
IIR	Infinite impulse response (filter)
Lin	Order of internal component of spherical expansion (Maxwell filter)
Lout	Order of external component of spherical expansion (Maxwell filter)
MNE	Minimum norm estimation
MPRAGE	Magnetisation prepared rapid gradient echo (MRI)
T1	Structural MRI scan
TA	Acquisition time (MRI)
TE	Echo time (MRI)
TR	Repetition time (MRI)
tSSS	Spatiotemporal signal space separation (Maxwell filter)

Physical Units

A*m	Ampere metre
AU	Arbitrary unit
dBFS	Decibels relative to full scale
fT/cm	Femtotesla per centimetre
Hz	Hertz
kHz	Kilohertz
S/m	Siemens per metre

General Research Terms and Notation

3D	Three-dimensional
Clu	Cluster
H.ID	Hypothesis ID
H.ID*e	Exploratory hypothesis ID
I	Instruction
ITI	Inter-stimulus interval
OSF	Open Science Framework
Prob	Probability
R.	Research question
ROI	Region of interest
TOI	Time of interest
v.	Package version

List of Figures

1.1 Overview of Key Components from Cognitive and Affective Neuropsychological Frameworks . 3
2.1 Graphical Abstract 22
2.2 Brain Networks Involved in Facial Emotion Discrimination and Visuo-Spatial Working Memory . 24
2.3 Dual-Task Experiment With Facial Emotion Discrimination and Visuo-Spatial N-Back 30
2.4 Post-Hoc Comparisons of the Behavioural Main Effects 39
2.5 Spatiotemporal Clusters for Significant Comparisons of Emotional Expression 41
2.6 Significant Alpha- and Low Beta- Band Cluster 43
2.7 Post-Hoc Comparisons of the Gaze-Related Main Effects 44
3.1 Bootstrapped Grand Averages for the Continuous Valence Ratings 77
3.2 Bootstrapped Grand Averages for the Continuous Arousal Ratings 78
3.3 Bootstrapped Average Valence, Arousal, and Dominance Ratings 79
3.4 Bootstrapped Average Standard Deviation of the Valence and Arousal Ratings 80
3.5 Distribution of the Clustered Data on the Three SAM Scales 81
3.6 Repeated Pairwise Subtractive Comparisons 82
3.7 Bootstrapped Average Familiarity Ratings 84
3.8 Frequencies of the Basic and GEW Emotions Ratings 84
3.9 Bootstrapped Average of the Basic and GEW Emotions Ratings 85
4.1 Experimental Set-Up of the Driving Simulator 101
4.2 Post-Hoc Comparisons of Interaction Effects in the Subjective Ratings 111
4.3 Post-Hoc Comparisons of Main Effects in the Subjective Ratings 113
4.4 Post-Hoc Comparisons of Main Effects in the Gaze Behaviour Measures 114
4.5 Results of the Permutation-Based Spatial Clustering of Source Power Localised with DICS . . 116
4.6 Average Performance and Confusion Matrix of the Decoding 118
4.7 Three-way Interaction Between Decoding Group, Valence of Speech, and Workload 119
4.8 Decoding Patterns of the Four Common Spatial Pattern Components 120

5.1 Visual Summary of the Key Findings from the MEG–eye-tracking Study I and Study III 143

6.1 Overview of the Experiment Procedure 176
6.2 Spatiotemporal *F*-Statistic Cluster of the Main Effects on Event-Related Magnetic Fields 178
6.3 Regions of Interest Comprising the Primary and Secondary Peak 180
6.4 Correlation Matrices of the Correlations Between Regions of Interest 181

6.5 Correlation Between Load-Based Alpha Band Modulation and N-Back Retrieval Reaction Time 183
6.6 Defined Regions of Interest for the Fixation-Related Eye-Tracking Analysis 186
6.7 Exploratory Interaction Comparisons of Task Phase and Face Region 187
6.8 Descriptive Heatmap of the Fixations During the Facial Emotion Discrimination 188
6.9 Descriptive Heatmap of the Fixations During the N-Back Encoding/Retrieval Phase 189

7.1 Results of the Personality Questionnaire Big Five 197
7.2 Results in Questionnaires Associated With Subjective Well-Being and Anxiety 198
7.3 Illustration of the Presented SAM and GEW Scales 202
7.4 Bootstrapped Mean of the Probability of Cluster Affiliation per Cluster for Critical Sequences . . 208
7.5 Frequencies of the Basic Emotions Ratings for the Negative Audio Sequences 213
7.6 Frequencies of the Basic Emotions Ratings for the Neutral Audio Sequences 214
7.7 Frequencies of the Basic Emotions Ratings for the Positive Audio Sequences 215
7.8 Frequencies of the GEW Emotions Ratings for the Negative Audio Sequences 216
7.9 Frequencies of the GEW Emotions Ratings for the Neutral Audio Sequences 217
7.10 Frequencies of the GEW Emotions Ratings for the Positive Audio Sequences 218
7.11 Average Cohen's d as Measure of Discriminatory Power Relative to the Neutral Condition . . . 226

8.1 Visualisation of the Steps Involved in Computing the Index of Pupillary Activity 239
8.2 Descriptive Statistics for Subjective Ratings 240
8.3 Descriptive Statistics for Correlates of Gaze and Driving Behaviour 240
8.4 Frequency Spectra of Source-Localised Oscillatory Power Obtained Using DICS 241
8.5 Oscillatory Source Power After Subtraction of the Aperiodic Component in the Main Effects . . 242
8.6 Oscillatory Source Power Without Subtraction of the Aperiodic Component 243
8.7 Descriptive Statistics for Source Power in Frequency Bands Within Significant Clusters 244
8.8 Spatial Clustering of the Main Effect Emotional Speech on the Aperiodic Components 244
8.9 Spatial Clustering of the Main Emotion Effect Without Subtracting the Aperiodic Component . . 245
8.10 Bootstrapped Mean F1 Scores Across Folds per Subject 246
8.11 Average Gamma Source Power Across Vertices From the Decoding Components 247
8.12 Significant Clusters of the Gamma-Band Interaction Effect and Source Estimates 248
8.13 Spatial Clustering of the Interaction Effect on the Aperiodic Broadband Component 249
8.14 Significant Clusters of Gamma-Band Interaction Without Subtracting the Aperiodic Component 250

List of Tables

2.1 Summary of the Hypotheses 28

6.1 Summary of the Behavioural Linear Mixed-Effects Models 177
6.2 Descriptive Parameters of the Regions of Interest 179
6.3 Correlation Between Alpha and Low Beta Cluster Power and N-Back Performance 183
6.4 Summary of the Gaze-Related Linear Mixed-Effects Models 185

7.2 Bootstrapped Mean of the Big Five, Subjective Well-Being, and Anxiety Questionnaires 199
7.3 Internal Consistency for the Used Personality, Well-Being, and Anxiety Scales 199
7.4 Bootstrapped Average Valence, Arousal, and Dominance Ratings per Audio Sequence 203
7.5 Bootstrapped Average Standard Deviation of the Valence and Arousal Ratings 206
7.6 Probabilities for Assignment to Negative, Neutral, and Positive Clusters per Sequence 209
7.7 Bootstrapped Average Familiarity Ratings per Audio Sequence 211
7.8 Pearson Correlations of the SAM Ratings Valence, Arousal, and Dominance With Familiarity . 212
7.9 Bootstrapped Average Negative and Positive Basic Emotions per Audio Sequence 219
7.10 Bootstrapped Average Negative and Positive GEW Emotions per Audio Sequence 221
7.11 Bootstrapped GEW Emotion Intensity Ratings per Audio Sequence 223
7.12 Results of the Repeated Measure Analyses of Variance With Condition as Within-Subject Factor 224

8.1 Descriptive Statistics of Centre Frequencies per Power Band in Hz 231
8.2 Summary of the Linear Mixed-Effects Models for Subjective Ratings 233
8.3 Summary of the Linear Mixed-Effects Models for Supplementary Subjective Ratings 235
8.4 Summary of the Linear Mixed-Effects Models for Driving Quality and Gaze Behaviour Measures 236

1 General Introduction

The human brain has evolved to be intrinsically social, enabling both automatic and flexible decision-making and behaviour in response to others (Adolphs, 2009; Frith, 2007). Displaying and interpreting emotional expressions play a key role in navigating social interaction. Such emotional cues allow us to convey and infer intentions, feelings, and thoughts, in order to achieve a personal or shared goal (Adolphs, 2009).

Hence, the processing and interpretation of emotional information are intricately linked to goal-oriented behaviour. Emotional cues can facilitate sensory processing (Schupp et al., 2003; Zeelenberg et al., 2006), attention allocation (Phelps et al., 2006; Vuilleumier, 2005), executive functioning (Lindström & Bohlin, 2011), and decision-making (Schwarz, 2000), when they align with the goal (see Schweizer et al., 2019, for meta-analytic review). However, when they are irrelevant to the primary goal, they still capture attention (Carretié, 2014) and divert cognitive resources away from goal-oriented processes (see Dolcos et al., 2011; Schweizer et al., 2019, for reviews). Specifically, emotional faces tend to persistently engage attentional and cognitive resources, even after they have been thoroughly processed (E. Fox et al., 2002; Sheppes et al., 2015). In such scenarios, the ability to control and modulate emotional processing is vital to minimise goal interference and prevent impaired behaviour arising from emotional distraction (Dolcos et al., 2011; J. J. Gross, 2002, 2015a; J. Zhang et al., 2023).

Initially, in psychology and neuroscience, emotional and cognitive processes were studied in isolation, both conceptually and neurophysiologically. However, over the past two decades, a paradigm shift has occurred towards recognising the interdependent nature of these processes, resulting in a growing body of research investigating their interaction and integration (see Dolcos et al., 2011; Ochsner et al., 2012; Pessoa, 2008, for reviews). Drawing on evidence from neuroimaging research and network modelling, influential reviews argue that the neural circuits underlying cognitive control, primarily involving prefrontal areas, communicate closely with subcortical regions implicated in affective processing and emotional responses (Dolcos et al., 2011; Gray et al., 2002; Pessoa, 2008). The authors propose a bilateral influence between emotional and cognitive processes and conceptualise the communication between the involved neural circuits and, to some degree, the integration into a shared neural network.

To conclude, it is now well established that emotional stimuli influence cognitive processes, such as memory, attention, and executive function, while cognitive states, such as depleting cognitive resources during increased load, in turn, modulate emotional processing (Dolcos et al., 2011).

Remarkably little is known about how the valence of an emotional stimulus and the current level of cognitive load co-modulate neural processing; despite evidence on differences regarding claimed attentional and processing resources (Becker & Rheem, 2020; Calvo & Beltrán, 2013; but cf. Savage et al., 2013), lateralisation (Davidson, 1992; Fusar-Poli et al., 2009; Grimshaw & Carmel, 2014; Harmon-Jones et al., 2010),

and even engaged brain regions (Kensinger & Schacter, 2006) between emotionally positive and negative stimuli. This is primarily because most studies have focused either exclusively on negative emotions or have examined only a single level of cognitive load (Ochsner et al., 2012; Schweizer et al., 2019).

In sum, a comprehensive understanding of *how* and *when* emotional and cognitive processes co-modulate neural information processing and control mechanisms remains incomplete (Brockhoff et al., 2022; Dolcos et al., 2011). This is particularly the case in dual-task paradigms where emotional information is concurrently or temporarily task-relevant (hereafter referred to as emotion-cognition dual tasks).

The following Section 1.1 presents theoretical frameworks from affective and cognitive psychology and neuroscience that help to conceptualise mechanisms underlying the interaction between emotional and cognitive processes. Section 1.2 then reviews the literature on functional brain networks and neural mechanisms implicated in emotion-cognition interactions in dual-task scenarios. Sections 1.3 and 1.4 introduce methodological approaches and experimental paradigms suited to examining this interplay. Finally, Section 1.5 outlines the overall scope and objectives of the thesis, presenting the central research questions and empirical studies that address them.

1.1 The Theoretical Basis of Interacting Emotional and Cognitive Processes

Being at the intersection of two research fields and communities, there is currently no integrative theoretical framework that explains when and how emotional and cognitive processes interact in dual-task paradigms. However, a plethora of theories from cognitive and affective neuropsychology offer conceptualisation of particular subcomponents involved in the interaction of emotional and cognitive processes. These subcomponents include attention allocation (see Oberauer, 2019; Wickens, 2021, for reviews), limited capacity of cognitive resources (Baddeley, 1992; Cowan, 1999; Wickens, 2002), information processing and cognitive control under different perceptual and working memory (WM) load levels (Lavie, 2010; Murphy et al., 2016), emotion appraisal (Scherer et al., 2001) and emotion regulation (J. J. Gross, 2015b; Ochsner & Gross, 2005; Ochsner et al., 2012).

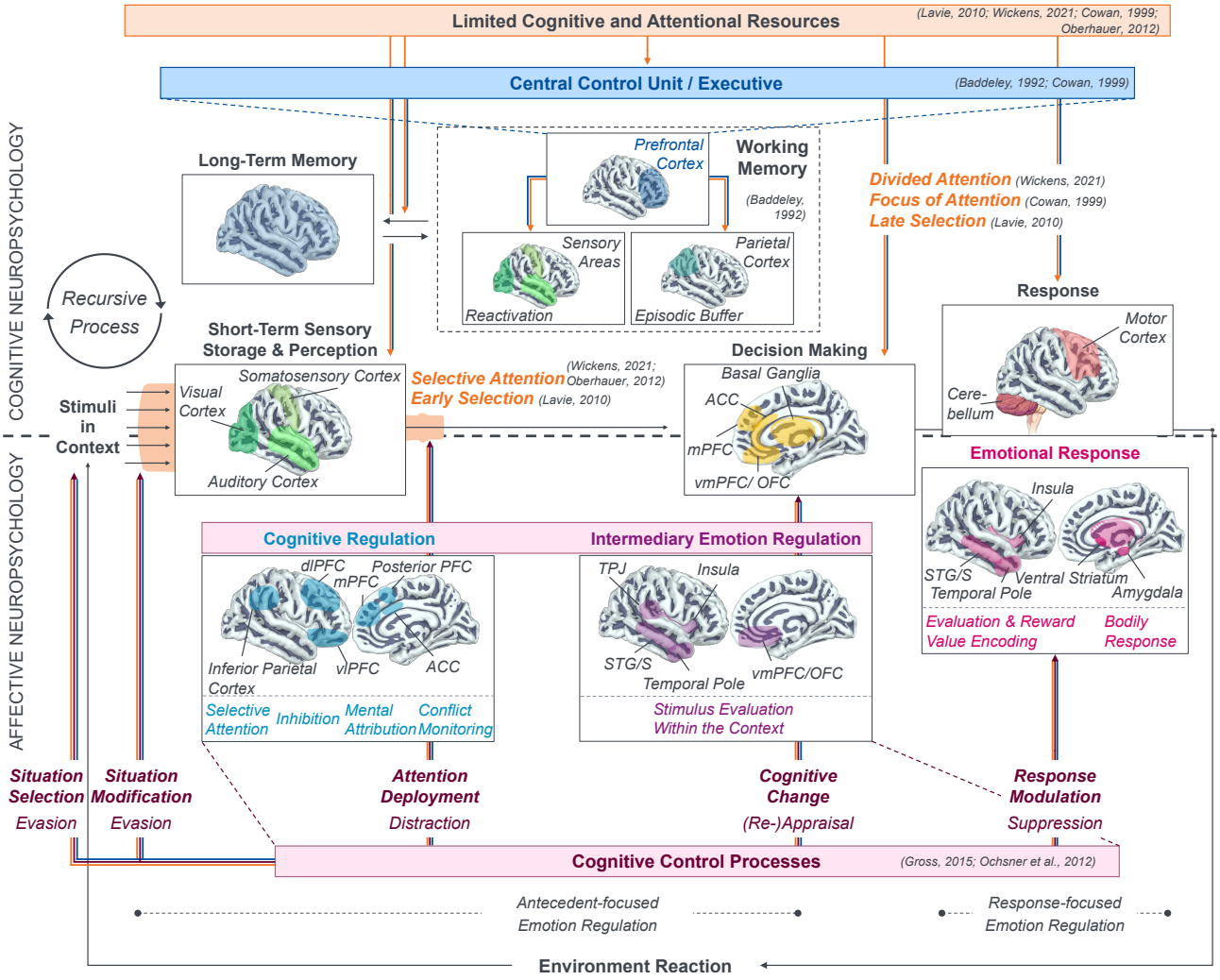
Figure 1.1 presents an alignment and integration of these subcomponents, together with the relevant brain regions involved, for studying interacting emotional and cognitive processes.

1.1.1 Limited Cognitive and Attentional Resources

A core assumption when investigating the interaction of emotional and cognitive processes is that cognitive resources are limited, including the capacity to control attention allocation, as well as process and retain information (Figure 1.1). This assumption is articulated more explicitly in models from cognitive neuropsychology (Cowan, 1999; Lavie, 2010; Oberauer, 2019; Wickens, 2021, upper part of Figure 1.1), than in those from affective neuropsychology (J. J. Gross, 2015b; Ochsner & Gross, 2005; Ochsner et al., 2012, lower part of Figure 1.1).

Figure 1.1

Schematic Overview of Key Components from Cognitive and Affective Neuropsychological Frameworks Implicated in the Interaction Between Emotional and Cognitive Processes



Note. The figure integrates principal constructs and theories from cognitive neuropsychology (upper part) and affective neuropsychology (lower part), including attention allocation (orange and blue components), information processing, and regulatory control over emotional processes (pink components). Moving from left to right in a clockwise direction, it illustrates the recursive sequence of processing (emotional) stimuli and responding to them within the context of the current environment. Attentional, cognitive, and regulatory mechanisms, along with information from long-term and working memory, can modulate each stage of processing, including perception in sensory brain regions, evaluation, decision-making, and the resulting (emotional) response. Up to the final stage of an elicited response, regulatory strategies are antecedent-focused; afterwards, they shift to being response-focused, such as expressive suppression (J. J. Gross, 2015a). Cognitive and attentional resources, illustrated by orange arrows, are limited and influence both executive control (blue arrows) and regulatory control (pink arrows). Multi-coloured arrows visualise this joint effect. ACC: anterior cingulate cortex; dlPFC: dorsolateral prefrontal cortex; mPFC: medial prefrontal cortex; OFC: orbitofrontal cortex; STG: superior temporal gyrus; STS: superior temporal sulcus; TPJ: temporo-parietal junction; vlPFC: ventrolateral prefrontal cortex; vmPFC: ventromedial prefrontal cortex. Adapted from Baddeley (1992), Cowan (1999), J. J. Gross (1998), and Wickens (2021). Brain images illustrating cognitive control processes are adapted from Ochsner et al. (2012).

Attentional resources guide perception, processing, and response selection through two mechanisms (see Oberauer, 2019; Wickens, 2021, for reviews): (1) Selective attention operating as a filter for incoming information (Driver, 2001; Murphy et al., 2016), and (2) divided attention allocating resources simultaneously across multiple components along the whole processing stream (e.g., perceiving new information while maintaining other task-relevant information in WM; see orange-blue arrows in Figure 1.1; Cowan, 1999; Oberauer, 2019; Wickens, 2021). The prefrontal cortex (PFC) acts as a central control unit for executive functioning, directing the allocation of attentional and cognitive resources towards goal-relevant processing (Baddeley, 1992; Cowan, 1999; Friedman & Robbins, 2022; Jurado & Rosselli, 2007; Koechlin et al., 2003, Figure 1.1, blue component). The WM system comprises the prefrontal control unit, as well as parietal and sensory regions, to temporarily maintain and manipulate task-relevant information (Baddeley, 1992; Cowan, 1999; Jurado & Rosselli, 2007). Importantly, attentional resources can be allocated inappropriately (e.g., to an emotional distractor) or insufficiently to a task-relevant process (e.g., maintaining information in WM). This leads to performance degradation, including the decay of stored information (Cowan, 1999; Oberauer, 2019).

1.1.2 Task Demands and Cognitive Load

The amount of information that must be stored and processed in WM simultaneously is defined as WM load (Baddeley, 1992; Cowan, 1999). WM capacity refers to the individual limit for maintaining and manipulating items in WM (Oberauer, 2019). The definitions of cognitive load (most frequently used in the field of educational psychology; Sweller et al., 1998) and mental workload (a broader concept from the fields of ergonomics and human factors; Young et al., 2015) both incorporate individual WM capacity in determining the current experienced load. The experienced cognitive (or mental) load reflects the proportion of resources required to process information and perform a given task (task demands) relative to the available resources. Underload occurs when task demands are well below available resources, and overload arises when demands exceed them (Sweller et al., 1998; Young et al., 2015). The relationship between cognitive load and performance is non-linear, with peak performance at moderate load, and reduced performance at both very low and high load levels (Sweller et al., 1998; Young et al., 2015). Notably, the degree of experienced load further influences the subjective emotional experience (e.g., boredom during underload and frustration during overload; Baker et al., 2010; Inzlicht et al., 2015). Affective information, in turn, influences the level of experienced load during a task (e.g., Gendolla & Silvestrini, 2011; Grissmann et al., 2017; Lingelbach et al., 2023; J. Zhang et al., 2023).

Dual- and Multi-Tasking

During dual- and multi-tasking, when more than one task is performed simultaneously – as is common in complex environments – subprocesses must be coordinated (Borst et al., 2010; Salvucci & Taatgen, 2008) and distributed across available resources accordingly (Norman & Shallice, 1986; Pashler, 1994; Wickens, 2002). These subprocesses draw on processing and response resources to acquire information from and interact with the environment. The resources themselves are multidimensional, involving distinct perceptual (visual, auditory, and tactile) and response modalities (vocal and tactile; e.g., Salvucci & Taatgen, 2008;

Wickens, 2002).

The theory of threaded cognition by Salvucci and Taatgen (2008) offers a detailed cognitive modelling architecture that integrates key concepts from related theories (Norman & Shallice, 1986; Pashler, 1994; Wickens, 2002). It conceptualises the coordination and processing of multi-tasking subprocesses, referred to as threads and their associated resource requests, as domain-independent. The architecture encompasses both serial and parallel processing, temporal sequence predictions, as well as learning mechanisms and effects (Salvucci & Taatgen, 2008). It assumes serial, one-at-a-time processing of requests within each resource and parallel processing across distinct resources (see also Wickens, 2002). A central cognitive unit, referred to as the procedural resource, communicates with other resources and coordinates the scheduling and distribution of processing requests. Free resources are claimed by requests as soon as they are needed (greedy principle) and released as soon as they are no longer required (polite principle). This central unit also includes: (1) a buffer that stores overarching goals (akin to the concept of a WM but limited to one chunk; Baddeley, 1992; Cowan, 1999), and (2) a declarative resource, comparable to long-term memory, which stores static knowledge and enables learning (Salvucci & Taatgen, 2008).

At the level of this central unit, interference may occur within both the procedural resource (consistent with other theories that propose a central bottleneck at the executive level; Pashler, 1994) but also the declarative resource. In contrast to related theories, the model by Salvucci and Taatgen (2008) does not posit that the central cognitive unit functions as a supervisory attentional system (but cf. Norman & Shallice, 1986), nor does it assume that stored information in the buffers requires sustained attentional resources to prevent information decay (but cf. Cowan, 1999). Moreover, it omits explicit considerations of limited cognitive capacity constraints (but cf. Wickens, 2002). These different assumptions may oversimplify certain aspects of dual-tasking mechanisms and sources of interference – particularly at the neural level (Cowan, 1999; Oberauer, 2019) – in favour of more parsimonious cognitive modelling (Salvucci & Taatgen, 2008).

To conclude, dual-task interference arises when multiple time-critical subprocesses create a bottleneck by competing for the same perceptual, cognitive, or response resources (i.e., selective attention; Borst et al., 2010; Broadbent, 1958; Pashler, 1994; Salvucci & Taatgen, 2008; Wickens, 2002). Moreover, temporal overlap or proximity between subprocesses increases interference, as the resources required for one process remain engaged by the other (Duncan et al., 2021; Pashler, 1994; Salvucci & Taatgen, 2008). Interference may also occur when the combined task demands exceed available cognitive resources, and attention cannot be effectively distributed across all required subprocesses (i.e., capacity limits in divided attention; Cowan, 1999; Wickens, 2002). Collectively, these sources of interference result in performance degradation in at least one task under dual-task conditions (e.g., Salvucci & Taatgen, 2008; Wickens, 2002).

The following subsections elaborate on the role of emotional stimuli and processes, and their interaction with cognitive functions in dual tasks.

1.1.3 What are Emotions and Emotional Responses?

In light of the persistent debate regarding the definition of emotion in affective neuroscience and psychology (e.g., LeDoux, 2012; Russell, 1980; Scherer et al., 2001), I next clarify the characteristics of emotional stimuli, emotion perception, emotional responses, and differentiate emotion from affect and mood.

While internal events (e.g., thoughts and memories; K. C. Fox et al., 2018) also elicit an emotional response and experience, this thesis focuses on external socio-emotional stimuli, specifically facial expressions and naturalistic speech. Emotional facial expressions have been intensively studied to understand emotion processing (Fusar-Poli et al., 2009; Schirmer, 2018), attention to emotion (Schindler & Bublatzky, 2020; Weidner et al., 2022) and interactions with cognitive processes (Brockhoff et al., 2022). There is considerably less research on naturalistic emotional speech listening (Nummenmaa et al., 2014; Schirmer, 2018; Zioga et al., 2023; but see Chan & Singhal, 2015, for effects of single emotional words on dual-task processes). The limited availability of validated stimulus material partly explains this paucity of research employing naturalistic emotional speech.

Emotional facial expressions and speech are proposed to be inherently salient – a key criterion of an emotional stimulus (LeDoux, 2012) – due to their communicative relevance in social interactions (Adolphs, 2002, 2009), and fundamental role in individual survival and well-being (LeDoux, 2012).

Two dimensions characterise the emotional properties of a stimulus: its pleasantness or hedonic value (valence) and its potential for mobilisation (arousal; Bradley & Lang, 1994; Russell, 1980; Russell & Barrett, 1999; see Ekman, 1992b, for an alternative concept of discrete basic emotions). High arousal and intense valence, whether positive (high valence) or negative (low valence), facilitate the recognition of emotional expressions (Britton et al., 2006; Winston et al., 2003).

Notably, perceiving and recognising emotions expressed by others can elicit bodily precursors of an emotional response in the perceiver – a phenomenon referred to as embodied emotion (Niedenthal, 2007). This effect is particularly evident for high-arousing stimuli (Almeida et al., 2016; Kensinger & Schacter, 2006). However, the intensity and quality of the ultimately experienced emotion are determined not only by perceived stimulus characteristics but also by recurrent appraisal processes that evaluate its significance in light of the current context and individual goals. These evaluative processes further modulate elicited neurophysiological, endocrine, and bodily responses (LeDoux, 2012; Scherer et al., 2001). The evaluation supports cognitive functioning in the presence of emotional information and enables empathic, rather than emotionally contagious, responses to others' emotional states (J. J. Gross & Jazaieri, 2014; Lamm & Singer, 2010).

The term affect is generally less clearly defined (Alpert & Rosen, 1990). It is sometimes used interchangeably with emotion, at other times as an umbrella term encompassing both emotion and mood, or as a brief, arousing experience with less specific valence (Alpert & Rosen, 1990). In this thesis, I adopt the latter definition, conceptualising affect as a short-lived, arousing state. In contrast, emotion is characterised by more pronounced and thorough evaluation, as well as more differentiated valence (J. J. Gross, 2002, 2015a; Scherer et al., 2001). By contrast, moods typically persist longer than emotions and are not necessarily tied to one specific cause (J. J. Gross, 1998, 2015a).

To conclude, an emotional experience unfolds over an emotion-generative sequence of subprocesses

(including perception, evaluation, and response, LeDoux, 2012; Scherer et al., 2001, see also Figure 1.1). Emotional processes as well as their interaction with cognitive processes can be examined at the neural, behavioural, and subjective level (Braunstein et al., 2017; Ochsner et al., 2012).

1.1.4 Towards the Co-Modulation of Emotional Valence and Cognitive Load

To investigate how the valence of emotional information and cognitive load specifically co-modulate available resources, attentional control, and information processing, two theories offer strong explanatory power: (1) *the load theory of selective attention and cognitive control* (Lavie, 2005, 2010; Lavie et al., 2004), explaining how emotional distractors influence cognitive processes, and (2) *the process model of emotion regulation* developed by J. J. Gross (1998, 2002, 2015b) with its neural mechanisms outlined by Ochsner and Gross (2005) and Ochsner et al. (2012), explaining how cognitive processes influence emotion processing and regulation.

The Load Theory of Selective Attention and Cognitive Control

Based on behavioural, electrophysiological, and neuroimaging evidence (see Murphy et al., 2016, for review), this theory suggests that both the level of load (low vs. high) and type of load (perceptual vs. executive functioning/WM) modulate attention allocation towards goal-relevant stimuli and cognitive control in the presence of interfering distractors (Lavie, 2005, 2010; Lavie et al., 2004). The theory proposes that under high perceptual load, attention is selectively allocated to task-relevant information, resulting in reduced distractor interference because no resources remain available to process task-irrelevant information. During low perceptual load, spare cognitive resources allow the processing of task-irrelevant information, which results in detrimental interference effects (Lavie, 2010, load effects on selective attention in an early processing stage, Figure 1.1). Conversely, for load targeting the executive functioning and WM load, the authors propose the opposite relationship: less distractor interference occurs under low load, while increased distractor interference arises under high WM load (Lavie, 2010; Lavie et al., 2004, load effects on attentional resources and executive functioning during later stages, Figure 1.1).

Effects of Emotional Stimuli on Cognitive Processes. Originally developed for non-emotional distractors, the theory has been extended to account for the influence of emotional information on cognitive control and processing across varying load levels (see Brockhoff et al., 2022, for review). Evidence from electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) indicates that emotional processing and enhanced distraction by emotionally, compared to neutral, stimuli is pronounced during low, but not high, perceptual load (Brockhoff et al., 2022). However, relatively few studies have examined WM load (F. Cao et al., 2022; Tavares et al., 2016), included emotionally positive stimuli (Tavares et al., 2016; van Dillen & Derks, 2012), or presented emotional information auditorily (Lingelbach et al., 2023; Mothes-Lasch et al., 2012). Their findings suggest emotional and valence-specific processing during low WM load (F. Cao et al., 2022; van Dillen & Derks, 2012; but cf. Brockhoff et al., 2022; Lavie et al., 2004), and pronounced distractor effects during high

WM load (Schweizer et al., 2019; Tavares et al., 2016).

In summary, the theory explains how cognitive load affects attention allocation towards and the processing of emotional distractors. However, it does not address scenarios in which emotional information is temporally task-relevant or voluntarily attended, nor does it specify how the effect of emotional information is regulated to prevent interference. The regulation of emotional responses elicited by perceived stimuli is addressed in the following subsection, which introduces the process model of emotion regulation (Figure 1.1; lower part).

The Process Model of Emotion Regulation

The process model of emotion regulation delineates cognitive mechanisms that modulate naturally evolving responses elicited by an emotional stimulus or event (see also Figure 1.1; lower half and pink components; J. J. Gross, 1998, 2002, 2015b; J. J. Gross & Thompson, 2007).

Effects of Cognitive Processes on Emotion. The process model of emotion regulation posits that regulatory processes applied at various stages of the emotion-generative sequence alter the emotional experience (its nature, magnitude, and duration; Figure 1.1, lower part; see J. J. Gross, 1998; J. J. Gross & Thompson, 2007; Koole, 2010, for comprehensive reviews). The regulatory processes can target early emotion-generative processes (antecedent-focused regulation), or the final stage of an emotional response (response-focused regulation; J. J. Gross, 1998, 2015b). The model introduces the following regulatory strategies (Figure 1.1): situation selection (e.g., evasion before encountering the stimulus), situation modification (e.g., evasion after encountering the stimulus), attention deployment (e.g., through distraction), cognitive change (e.g., reappraisal), and response modulation (e.g., suppression). Emotion regulation comprises the identification of an emotional response precursor as a candidate for regulation (depending on context, current goals, and resources), selection of a regulatory strategy, implementation, and finally, monitoring with potential re-entry into the cycle (J. J. Gross, 2015b; Kalisch, 2009; Pruessner et al., 2020, see also recursive loop in Figure 1.1).

Importantly, parallels can be drawn to earlier discussed theories from cognitive neuropsychology: Emotion regulation is proposed to be achieved through cognitive control (Ochsner & Gross, 2005; Ochsner et al., 2012)[1]. Ochsner and Gross (2005) describe cognitive control as a broadly defined set of processes ranging from attentional control to cognitive change. Koechlin et al. (2003) define it more specifically as ”the ability to coordinate thoughts and actions in relation to internal goals” (p. 1181), emphasising the role of the lateral PFC in exerting cognitive control. In their cascade model of cognitive control, Koechlin et al. (2003) proposed that the lateral PFC is functionally organised in premotor regions involved in sensory control, caudal regions linked to contextual control, and rostral regions implicated in episodic control (Koechlin et al., 2003).

When integrating the concept of emotion regulation into a broader framework (Cowan, 1999; Oberauer, 2019; Wickens, 2021), one can summarise that it (1) draws on the available resources (referred to in the field of emotion regulation as regulatory or cognitive effort; Shepherd, 2023), (2) corresponds at least partly to

[1] The conceptual differentiation and relation between cognitive control and emotion regulation are relatively ambiguous (J. J. Gross, 2015a; Ochsner & Gross, 2005; Ochsner et al., 2012; Pruessner et al., 2020).

executive functioning (Koechlin et al., 2003; Pruessner et al., 2020), and (3) is, thus, closely linked to WM (Friedman & Miyake, 2017; Jurado & Rosselli, 2007; Pruessner et al., 2020). The amount of resources available for cognitive control also influences the effectiveness of the regulation (Wagner & Heatherton, 2012, on a neural level; Grillon et al., 2015, on a subjective level). Moreover, the initiation, maintenance, and strategy choice of emotion regulation appears to be guided by a cost-benefit analysis (Sheppes, 2020), with positive emotion reducing the cost (Higgins, 2006) and cognitive load increasing the cost associated with regulation (see Grahek et al., 2020, for a computational modelling approach).

In conclusion, the process model of emotion regulation offers a framework to understand how and when emotion-generative processes and emotional experiences are modulated by cognitive processes, depending on the available resources[2] and valence of a stimulus.

In the following section, I review neuroscientific evidence on key brain regions and networks involved in emotion-cognition interactions. I begin by describing the localisation of these regions and networks, followed by an outline of how neural rhythms coordinate information processing.

1.2 The Neurophysiological Basis of Interacting Emotional and Cognitive Processes

The brain networks involved in emotion-cognition interactions can be functionally categorised into: 1) emotion perception and recognition (emotional component); 2) executive functioning, attention, and WM under increased cognitive load (cognitive component); and 3) emotion regulation (integrative component). The following subsections are structured according to this functional categorisation of emotional, cognitive, and integrative components.

1.2.1 Neural Systems Involved in Emotion Perception

Meta-analytic evidence indicates modality-specific processing for emotional faces (Bernstein & Yovel, 2015; Fusar-Poli et al., 2009; Vuilleumier & Pourtois, 2007) and speech (Frühholz & Grandjean, 2013; Grandjean, 2021; Schirmer, 2018), with enhanced activation along the face- and speech-specific processing pathways when compared to non-emotional stimuli (Frühholz & Grandjean, 2013; Haxby et al., 2000; Vuilleumier & Pourtois, 2007).

Emotional facial stimuli are processed along a visual ventral stream, eliciting a particularly strong response in the fusiform gyrus of the inferior temporal cortex 170 ms after face onset (Bernstein & Yovel, 2015; Kanwisher et al., 1997; Kanwisher & Yovel, 2006). This response has been linked to a configural analysis of the facial expression (Bernstein & Yovel, 2015; Monroe et al., 2013; Schindler & Bublatzky, 2020).

Emotional speech encompasses both semantics (what is said) and paralinguistics (how it is said), with

[2] However, most existing studies focus primarily on explicitly instructed emotion regulation (Pruessner et al., 2020), without investigating contextual effects such as available resources and varying cognitive load levels.

prosody referring to the rhythmic melody (Belyk & Brown, 2014; Schirmer & Kotz, 2006). The superior temporal gyrus (STG) and sulcus (STS), along the auditory ventral stream, are particularly implicated in the listening to voices (also called temporal voice areas; Grandjean, 2021; Hickok & Poeppel, 2007; Schirmer, 2018).

A meta-analysis by Schirmer (2018) indicates that, compared with emotional voices, emotional faces more strongly engage the amygdala and medial temporal lobe. Importantly, perceiving both emotional faces and speech activates brain regions associated with social cognition and the mirror neuron system (Grandjean, 2021; Olson et al., 2007; Said et al., 2011; Scott, 2019). In addition, fronto-temporal regions, including the orbitofrontal cortex (OFC), anterior insula, and temporal pole (TP), are implicated in social decision-making, evaluation of socio-emotional information within the current context, and the binding of perceptual inputs to visceral emotional responses, respectively (Grandjean, 2021; Olson et al., 2007).

1.2.2 Neural Systems Involved in Cognitive Load

Increasing cognitive load in a task leads to an enhanced activation in regions involved in task-relevant components such as executive functioning, perception, WM, and response selection (see also upper part of Figure 1.1; and Rottschy et al., 2012, for meta-analysis).

In particular, prefrontal regions, including the lateral PFC and anterior cingulate cortex (ACC), are associated with executive functioning under varying cognitive load (Friedman & Robbins, 2022; Koechlin et al., 2003). These prefrontal regions are also closely connected to parietal areas, including the lateral parietal lobe and intraparietal sulcus, forming the fronto-parietal control network. This network is implicated in maintaining task performance through goal-directed stimulus-response selection under increased load (Corbetta & Shulman, 2002; Kim, 2019; Rottschy et al., 2012). During very high cognitive load, when resources are severely depleted, neural activation in task-relevant regions, such as sensory areas involved in perception and parietal regions involved in WM and spatial attention, tends to decrease (e.g., X. Chen et al., 2023; Michels et al., 2010).

1.2.3 Neural Systems Involved in Regulatory Control and Emotion Regulation

To coordinate and integrate interacting emotional and cognitive processes, Ochsner et al. (2012) proposed a distinction between two neural systems: one responsible for emotion regulation, and another serving as an intermediary between emotion generation and regulation (see Figure 1.1; pink-turquoise and pink-purple components, Ochsner & Gross, 2005; Ochsner et al., 2012). Both systems regulate emotion by mediating activity in subcortical emotion-generating regions, particularly the amygdala (Dolcos et al., 2011; Ochsner et al., 2012).

The first system encompasses the lateral PFC, posterior PFC (pPFC), and inferior parietal lobule (IPL; Ochsner et al., 2012). These regions, previously introduced in the context of executive functioning and goal-directed stimulus-response selection (Figure 1.1; blue component; Corbetta & Shulman, 2002; Friedman & Robbins, 2022; Koechlin et al., 2003), are proposed to initiate and maintain regulatory strategies, and, in the case of reappraisal, to manipulate the interpretation of emotional stimuli (J. J. Gross, 2002; Ochsner et al., 2012). The ACC is proposed to monitor the regulatory progress, while the dorsomedial PFC is implicated in

mentalising, that is, understanding one's own and others' emotional states, as well as thoughts, intentions and desires (Dixon et al., 2017; J. P. Mitchell, 2009; Ochsner et al., 2012; Olsson & Ochsner, 2008). The ability to represent one's own and others' mental states is known as Theory of Mind (ToM; Amodio & Frith, 2006; Olsson & Ochsner, 2008; Schlaffke et al., 2015; Schurz et al., 2021).

Ochsner et al. (2012) proposed that regulatory perceptual and semantic evaluations of emotional stimuli are orchestrated by the second system, which comprises the STG, STS, TP, ventromedial PFC (vmPFC), and regions around the temporo-parietal junction (TPJ).

Interestingly, Corbetta and Shulman (2002) previously proposed a similar functional distinction between a fronto-parietal control network, resembling Ochsner et al. (2012)'s first system, and a ventral parieto-temporal network, corresponding to their second system. The latter is thought to act as a neural "circuit breaker", interrupting ongoing cognitive activity to redirect attention towards salient environmental events (Corbetta & Shulman, 2002). The right TPJ, in particular, appears to support attentional reorienting (Krall et al., 2015). It is located at the intersection of the posterior STS and IPL, both associated with attention and spatial awareness (Corbetta & Shulman, 2002; Karnath, 2001).[3] The right TPJ facilitates mentalising and ToM through attentional reorienting and perspective-taking between self and others, thereby also enabling emotional distancing (Krall et al., 2015; van Overwalle, 2009).

Although not named by Ochsner et al. (2012), the insula and OFC are also likely to be involved in regulatory evaluations, particularly within a social context (Grandjean, 2021; Olson et al., 2007; Said et al., 2011; Scott, 2019).

Having outlined key brain regions and networks involved in emotion-cognition interactions, the focus now shifts to how neural rhythms shape information processing.

1.2.4 Information Processing Through Neural Oscillations

Synchronised firing of neural populations plays a fundamental role in the temporal coordination of neural processes (Buzsáki, 2006; Fries, 2005; Lopes da Silva, 2013; X.-J. Wang, 2010). Neural oscillations are commonly identified at specific frequencies, namely in the δ (0.2 - 3.5 Hz), θ (4 - 7.5 Hz), α (8 - 13 Hz), β (14 - 30 Hz), and γ bands ($>$ 30 Hz; Lopes da Silva, 2013). They serve as mechanisms for information coding, attentional modulation, and interareal communication (Buzsáki, 2006; Fries, 2005; Lopes da Silva, 2013; X.-J. Wang, 2010). Low-frequency oscillations ($<$ 14 Hz) are proposed to mediate large-scale communication across widespread brain regions, whereas high-frequency rhythms ($\geq$ 14 Hz) are associated with local neural processing (Engel et al., 2001; Lopes da Silva, 2013).

Numerous studies have reported decreased alpha band power and increased gamma band power during enhanced attention and sensory processing in cognitive tasks (Jensen & Mazaheri, 2010; see X.-J. Wang, 2010, for review). Alpha oscillations are implicated in regulating sensory processing via an inhibition-based gating mechanism, which modulates stimulus-induced gamma oscillations and cortical engagement (Bonnefond & Jensen, 2012; Jensen & Mazaheri, 2010; Jokisch & Jensen, 2007; Klimesch et al., 2007). They play a key

[3] Frequent co-activation of the right TPJ, STS, and IPL has sometimes obscured clear differentiation between these regions (Krall et al., 2015).

role in facilitating the processing of task-relevant information while inhibiting that of distractors (Bonnefond & Jensen, 2012; Jensen & Mazaheri, 2010). Gamma oscillations are proposed to contribute to sensory processing (Ou & Law, 2019), selective attention to relevant stimuli (Herrmann et al., 2010), and regulatory inhibition (Kang et al., 2014; Popov et al., 2012; Quirk & Beer, 2006). Notably, gamma activity positively correlates with the blood-oxygen-level-dependent (BOLD) responses measured by fMRI (Logothetis et al., 2001). In contrast, beta-band oscillations are proposed to maintain the "status quo": increased beta oscillations facilitate and stabilise ongoing cognitive and motor processes, whereas unexpected stimuli reduce beta-band power (Engel & Fries, 2010; Ou & Law, 2019). Beta oscillations are important for WM maintenance, motor control (Engel & Fries, 2010), and also speech tracking (Al-Zubaidi et al., 2025; Keitel et al., 2018; Ou & Law, 2019).

Magnetoencephalography (MEG) is well-suited for investigating the neural dynamics of information processing. It offers high temporal resolution (Lopes da Silva, 2013), while providing a spatial resolution in the millimetre range for cortical brain areas (J. Gross, 2019; Hämäläinen et al., 1993).

The following section describes MEG as a technique and outlines common approaches to analysing neural dynamics.

1.3 Magnetoencephalography

MEG is a non-invasive recording technique that measures extracranial magnetic fields generated primarily by postsynaptic currents in the dendrites of cortical pyramidal neurons (J. Gross, 2019; Lopes da Silva, 2013). The synchronous electrical activity in similarly oriented neuronal populations generates summated magnetic fields, which can be detected by the MEG sensors (Hämäläinen et al., 1993). MEG is most sensitive to tangential dipoles generated in the walls of cortical sulci, where apical dendrites of pyramidal neurons are oriented parallel to the scalp surface (reflecting changes in electromagnetic fields that are perpendicular to the sensors; Hämäläinen et al., 1993).

Conventional stationary whole-head MEG systems measure these summated, yet extremely weak, magnetic fields (in the range of femto-tesla) using approximately 300 sensors, which are Superconducting Quantum Interference Devices (SQUIDs; Cohen, 1972) coupled with flux transformers, consisting of a pickup coil(s) (Hämäläinen et al., 1993; Vrba & Robinson, 2001; see Boto et al., 2018, for advances in mobile MEG). These sensors are evenly distributed within a helmet-shaped dewar, which is a cryogenic vessel filled with liquid helium at around $-269\,^{\circ}\mathrm{C}$. This enables the sensors to operate in a superconducting state. When magnetic flux passes through the pickup coils, it induces a current proportional to the magnetic field strength. The pickup coils and SQUID sensors measure the changes in this current (Hämäläinen et al., 1993). To mitigate contamination from environmental magnetic fields, MEG systems are placed in magnetically shielded rooms.

Two types of pickup coils are commonly employed: (1) Magnetometers, which use a single coil to measure the magnetic field component perpendicular to the sensor surface; and (2) gradiometers, which measure the spatial gradient of the magnetic field (i.e., the difference between coils, expressed in femtotesla per centimetre; fT/cm). Magnetometers are sensitive to both nearby and distant sources but are prone to

external interference. Gradiometers, by contrast, are less affected by uniform external interference, as this is cancelled out when computing the spatial gradient. The latter can be arranged either as axial gradiometers, with coils aligned along the radial axis, or as planar gradiometers, with coils positioned side-by-side in the same plane. Axial gradiometers are sensitive to sources around the sensor, whereas planar gradiometers primarily detect sources directly beneath the sensor (tangential; Hämäläinen et al., 1993; Hansen et al., 2010).

Source space analysis of MEG data, which enhances spatial resolution, is introduced in the following subsection.

1.3.1 Source Space Analysis

Source reconstruction, also referred to as inverse modelling, estimates the locations and dynamics of the neural generators responsible for the magnetic fields recorded at the scalp (i.e., at the sensor level). Analyses performed in source space are motivated by the potential to increase spatial resolution by disentangling the contributions from different brain regions in the measured dynamics (J. Gross, 2019).

Two mathematical problems underlie neural source reconstruction: the forward and inverse problem. The forward problem entails computing a physical model that predicts the spatial distribution of the magnetic field at the sensors, given a set of neural current sources. In source reconstruction, these sources are typically modelled as current dipoles with defined position, orientation, and strength. Solving the forward problem for a given sensor configuration requires detailed knowledge of sensor properties and a head model that specifies the spatial distribution of tissue conductivities. Estimating tissue conductivity is more critical in EEG than in MEG, as electric field propagation is highly sensitive to conductivity, whereas magnetic fields are comparatively less affected. Two numerical methods, informed by individual brain anatomy from MRI scans, are commonly used to construct head models for source reconstruction: boundary element models (BEMs), which focus on segregated tissue surfaces (brain, skull, and scalp); and finite element models (FEMs), which model the conductivity of tissue throughout the entire head volume (Hämäläinen et al., 1993).

The inverse problem is ill-posed, meaning that many different current distributions can explain the data (non-uniqueness), and the solutions are sensitive to noise (ill-conditioning; Hämäläinen et al., 1993). To solve the inverse problem, various approaches estimate the location and orientation of current sources by introducing simplifying assumptions, approximations, and regularisation techniques, such as restricting sources to grey matter or constraining their orientation. Minimum norm estimation (MNE) and beamformers are two widely used inverse methods for MEG source localisation (Hansen et al., 2010). MNE is a distributed source modelling technique that estimates the most likely distribution of neural current sources by selecting the solution with the smallest overall L2 norm (i.e., the minimum total current amplitude) that still explains the observed magnetic fields. Beamformers, in contrast, construct spatial filters in source space to estimate activity from a target location, while suppressing contributions from other sources (similar to a beam that is directed exclusively at some region; Hansen et al., 2010).

The following subsections present methods for how to statistically analyse experimental effects in MEG data.

1.3.2 Statistical Analysis Using Permutation-Based Clustering

Notably, due to the spatiotemporal multidimensionality, MEG source space analyses face the statistical problem of multiple comparisons. Performing a statistical test at each observation unit (vertex-time point pair) inflates the probability of falsely finding a significant effect under the hypothesis of no effect (false positive or Type I error rate). Univariate approaches, such as an analysis of variance (ANOVA) or a *t*-test, address this by aggregating over a predefined region and time interval. An alternative approach is mass-univariate statistics, such as permutation-based clustering. This method allows data-driven statistical testing at every location and time point, while controlling for multiple comparisons (Maris & Oostenveld, 2007; see Groppe et al., 2011, for a tutorial review). Permutation-based clustering controls the Type I error rate by grouping time points and measurement positions (vertices in the case of MEG) into clusters based on temporal and spatial adjacency, thereby reducing the number of comparisons to the cluster level (Maris & Oostenveld, 2007).

To detect meaningful differences between conditions, a significance threshold is defined (e.g., $p < 0.05$). Condition labels are then randomly reassigned across clusters in thousands of permutations to generate a randomised dataset. A univariate statistical test (*F*- or *t*-test) is computed for each permutation, and a cluster is considered significant if the sum of the resulting statistical indices (*F*- or *t*-values) in the original data exceeds the 95^{th} percentile ($p < 0.05$) of the statistical value distribution obtained from the randomised data (controlling the Type I error rate; Maris & Oostenveld, 2007).

Although mass-univariate statistics may sacrifice statistical power compared to a univariate test (Groppe et al., 2011), it enhances statistical relevance for studies investigating temporally or spatially less well-understood phenomena – such is the case when investigating interacting cognitive processes or effects of complex, naturalistic stimuli – where latencies of evoked responses and spatial locations of interest cannot be easily predetermined (Groppe et al., 2011).

Permutation-based clustering has two attributes that can be disadvantageous in certain contexts: 1) it still analyses activation from clusters separately (Haufe et al., 2014), and 2) it assumes that neural activation patterns are anatomically and functionally similarly localised across subjects (Holdgraf et al., 2017; Marsicano et al., 2024). Consequently, this approach may be substantially limited when investigating complex neurocognitive dynamics involving individual-specific processing strategies (Marsicano et al., 2024), or naturalistic stimulus material (Sonkusare et al., 2019).

1.3.3 Statistical Analysis Using Multivariate Pattern Decoding

Multivariate pattern analysis (MVPA) increases sensitivity by integrating measurement sources within a multidimensional framework, while accounting for inter-individual neural variability (Haufe et al., 2014; Holdgraf et al., 2017; Marsicano et al., 2024). To achieve this, MVPA employs supervised machine learning (ML) techniques. When using a linear ML model, the model can be combined with inverse mapping methods to derive physiologically interpretable decoding patterns that differentiate between experimental conditions (Haufe et al., 2014; Holdgraf et al., 2017).

1.4 Studying the Emotion-Cognition Interaction in Naturalistic Contexts

To understand how the human brain processes information, researchers have systematically isolated and investigated single characteristics of highly controlled artificial stimuli. This approach ensures that observed phenomena can be directly attributed to the manipulated independent variables. However, a significant concern is that findings from such strictly controlled paradigms may not generalise to complex real-world scenarios, where stimuli are dynamic, perceptually rich, and embedded in a meaningful context (Fan et al., 2021; Hamilton & Huth, 2020; Redcay & Moraczewski, 2020). Notably, neurophysiological responses to more naturalistic stimuli often differ from those elicited by artificially fragmented stimuli, particularly in the domains of emotion processing (Goldberg et al., 2014; Xu et al., 2021), speech perception (Hamilton & Huth, 2020), and social cognition (Fan et al., 2021; Redcay & Moraczewski, 2020).

When moving from controlled to more complex, naturalistic research paradigms, two strategies can help derive meaningful insights: (1) richly annotating and validating the stimulus material (e.g., Häusler & Hanke, 2021; Labs et al., 2015), and (2) combining multiple measurement modalities (i.e., a multimodal measurement approach), such as neurophysiological, behavioural, and subjective measures, to achieve a comprehensive understanding and obtain converging evidence (e.g., Braunstein et al., 2017).

It is particularly advantageous to combine neurophysiological measures with eye-tracking, as this facilitates linking neural signatures to both the focus and content of visual attention and processing. Moreover, gaze and pupillary responses have been identified as reliable markers of information processing under different levels of cognitive load and emotional valence (see Skaramagkas et al., 2021, for comprehensive review), as well as during regulatory processes (e.g., Ferencová et al., 2021; Kinner et al., 2017; van der Wel & van Steenbergen, 2018).

1.5 Scope and Objectives of the Thesis

The study of cognitive and emotional processes was long regarded as two independent research fields in psychology and neuroscience. However, recent work increasingly recognises their interdependence and mutual influence, particularly in social contexts. Two existing frameworks provide a theoretical basis for aspects of the interaction between cognitive and emotional processes: The load theory of selective attention and cognitive control (Lavie, 2010) sheds light on how emotional distractors influence cognitive processing, while the process model of emotion regulation (J. J. Gross, 2015b; Ochsner et al., 2012) describes how regulatory mechanisms shape emotion processing. Despite this progress, the neurophysiological basis of the interplay between emotion and cognition remains not fully understood. More specifically, two key research questions have received very little attention:

- *Research Question I:* How are emotional and cognitive processes integrated in emotion-cognition dual-task paradigms where information in emotional stimuli is temporarily task-relevant?
- *Research Question II:* Under what circumstances do the valence of an emotional stimulus and the current level of cognitive load jointly influence information processing and regulatory control?

Insights into these open questions are highly relevant in situations where safety depends on effective performance, such as in the transportation sector, medical occupations, and first responder operations. Furthermore, they are crucial in clinical contexts: Various psychopathological disorders exhibit altered emotion processing and impaired integration of emotional information into cognitive functioning. These conditions are characterised by difficulties in expressing one's own emotions (e.g., autism spectrum disorder and alexithymia), interpreting others' emotional states and intentions (e.g., schizophrenia and depression), and executive functioning in the presence of affective information (e.g., attention deficit hyperactivity disorder, post-traumatic stress disorder, and anxiety; see Aldao et al., 2010; J. J. Gross & Jazaieri, 2014; Lincoln et al., 2022; Sheppes et al., 2015, for reviews). Understanding the neurophysiological basis of emotion-cognition interactions informs the development of interventions and adaptive assistive technologies designed to enhance well-being, social interaction, and performance in both healthy individuals and those with specific needs.

This thesis addresses the two outlined questions in two research projects comprising three studies. New insights are interpreted in the light and integrated into existing theoretical frameworks. The present work aims to provide empirical evidence that bridges the gap between different theoretical perspectives. However, it does not seek to validate any particular theory, model, or component of a model.

1.5.1 Emotional Face Perception and Visuo-Spatial Working Memory Load

The first project and study, introduced in chapter 2, utilised a dual task and MEG combined with eye-tracking to investigate emotional face processing, WM load, and the regulatory mechanisms involved in the interaction of emotional and cognitive processes. The study temporally sequentially interleaved the emotional and cognitive tasks. During an emotion discrimination task, naturalistic facial images, displaying negative, positive or neutral expressions, were presented as targets. Meanwhile, participants had to maintain information relevant to a visuo-spatial n-back task (i.e., the cognitive task) in WM. During the WM encoding and retrieval phase that followed, the faces acted as distractors. The visuo-spatial n-back task was performed either at low or high WM load. It was hypothesised that emotion-specific processing, that is, differences between facial valence conditions, would occur only under low, but not high, WM load in the perceptual task of emotion discrimination. However, under high, but not low, WM load, negative emotional expressions were expected to cause interference with the cognitive task, indicating diminished cognitive resources available to regulate their distracting effects. Permutation-based clustering of event-related magnetic fields (ERFs) and oscillatory activity in source space was employed to examine differences in neural processing between the experimental conditions. In addition, changes in fixation metrics and evoked pupil size were analysed. The study contributes to our understanding of a) when emotion-specific processing occurs, b) how regulatory processes facilitate disentanglement from the emotional component and enable an effective switch to the cognitive task, and c) how the valence of emotional faces and level of WM load influence available cognitive resources, information processing, and the likelihood of dual-task interference. The study was published in *Human Brain Mapping*.

1.5.2 Emotional Speech Perception and Visual-Spatial Workload

The second project examined information processing during the dual task of navigating a car through simulated driving scenarios at low and high visuo-spatial workload, while concurrently listening to emotional speech.

Development and Validation of a Naturalistic Speech Database for Emotion Induction

As there were no thoroughly validated German databases of naturalistic emotional speech available at the time, the second study, reported in chapter 3, developed and validated a novel auditory speech database specifically for emotion induction. Thirty-seven audio sequences, with a total duration of 92 minutes, were validated to elicit positive emotions using excerpts from comedian shows, neutral emotions with weather forecasts, and negative emotions through arguments between couples and relatives taken from films or television series. The key strength of the database is its rich annotation, which includes continuous ratings capturing the time course and variability of perceived valence and arousal on a per-second basis, as well as multiple discrete ratings (e.g., discrete emotions elicited, emotion intensity, and familiarity). Clustering and simulation analyses combined with bootstrapping quantified the extent to which the audio sequences met quality criteria for differentiation, salience and strength, as well as generalisability across participants. The study contributes a validated stimulus database of naturalistic speech suitable for investigating the processing and generation of emotions in German-speaking participants. The database is openly accessible on the Open Science Framework, and the study was published in *Behavior Research Methods*.

Studying the Emotion-Cognition Interaction in a Near-Realistic Multisensory Scenario

The third study, presented in chapter 4, utilised MEG, eye-tracking, and the validated stimuli to investigate multisensory information processing during driving while listening to emotionally negative, neutral, and positive speech. The driving task simulated executive functioning, spatial cognition, and WM in common real-world scenarios at low and high workload. Distinct gaze-related and oscillatory processing modes were expected for emotional speech tracking and visuo-spatial cognition at increased workload. Furthermore, regulatory mechanisms in the form of fronto-temporal oscillatory gamma activity were assumed to be co-modulated by both the valence of emotional speech and the current visuo-spatial workload. Hypotheses on neural processing modes were tested using permutation-based clustering of modulated oscillatory power in source space, as well as MVPA employing common spatial patterns combined with a linear discriminant analysis. The study provides converging evidence of distinct information processing modes for auditory emotional speech processing, visuo-spatial cognition, and control mechanisms that flexibly up- and down-regulate emotional interference in complex, naturalistic environments. It further advances the understanding of how emotional speech valence and visuo-spatial workload jointly influence cognitive resources and regulatory mechanisms. The study was under review and is now published in *Imaging Neuroscience*.

2 Study I – Neurophysiological Basis of Emotional Face Perception and Working Memory Load in a Dual-Task MEG Study

Katharina Lingelbach[1,2] and Jochem W. Rieger[1]

[1] Department of Psychology, University of Oldenburg, Oldenburg, Germany

[2] Fraunhofer Institute for Industrial Engineering IAO, Nobelstraße 12, 70569 Stuttgart, Germany

This article has been published as:

Lingelbach, K., & Rieger, J. W. (2025). Neurophysiological basis of emotional face perception and working memory load in a dual-task MEG study. *Human Brain Mapping, 46*(8), 1–25. https://doi.org/10.1002/hbm.70242

Author contributions:

Katharina Lingelbach: Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualisation, Writing – original draft, Writing – review & editing. **Jochem W. Rieger**: Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.

The original publication and this chapter are identical in content.

I hereby confirm that Katharina Lingelbach contributed to the aforementioned study as stated above.



Prof. Dr. Jochem W. Rieger

Oldenburg, 30.01.2026

Abstract

Research on the neurophysiological effects of emotional face processing, working memory (WM) load, and their interaction in dual tasks remains scarce. Therefore, we conducted a combined magnetoencephalography eye-tracking study with 47 participants. The dual task temporally interleaved a facial emotion discrimination task with a visuo-spatial n-back task. Source-space cluster analyses of event-related magnetic fields (ERFs) and oscillations revealed significant main effects of emotional expression and WM load. During emotion discrimination, enhanced ERFs for negative facial expressions located across the insula, anterior cingulate cortex, and face-specific occipital regions suggest amplified emotion processing but also the recruitment of attentional control mechanisms. During the n-back phase, emotional faces did not affect evoked responses when they were task-irrelevant. Interaction trends in pupil dilation indicated that emotion-specific processing is diminished under high WM load. During the n-back phase, increased WM load reduced alpha and low beta oscillations in temporo- and parieto-occipital areas. In addition, reduced target fixations in the presence of negative facial distractors indicated a tendency towards emotion-specific interference. Furthermore, sustained increased WM load affected perceived valence, pupil size, and reaction time in both subtasks. A convergence of neurophysiological, physiological, and behavioural findings points to specific processing modes with greater resource depletion for negative expressions and high WM load in the dual task. In conclusion, the study advanced our understanding of a) circumstances under which emotional faces modulate ERFs in a dual task, b) mechanisms underlying emotion discrimination, c) interaction effects of emotional expression and WM load in gaze behaviour, as well as d) how WM-related oscillatory alpha and beta power is affected by increasing load.

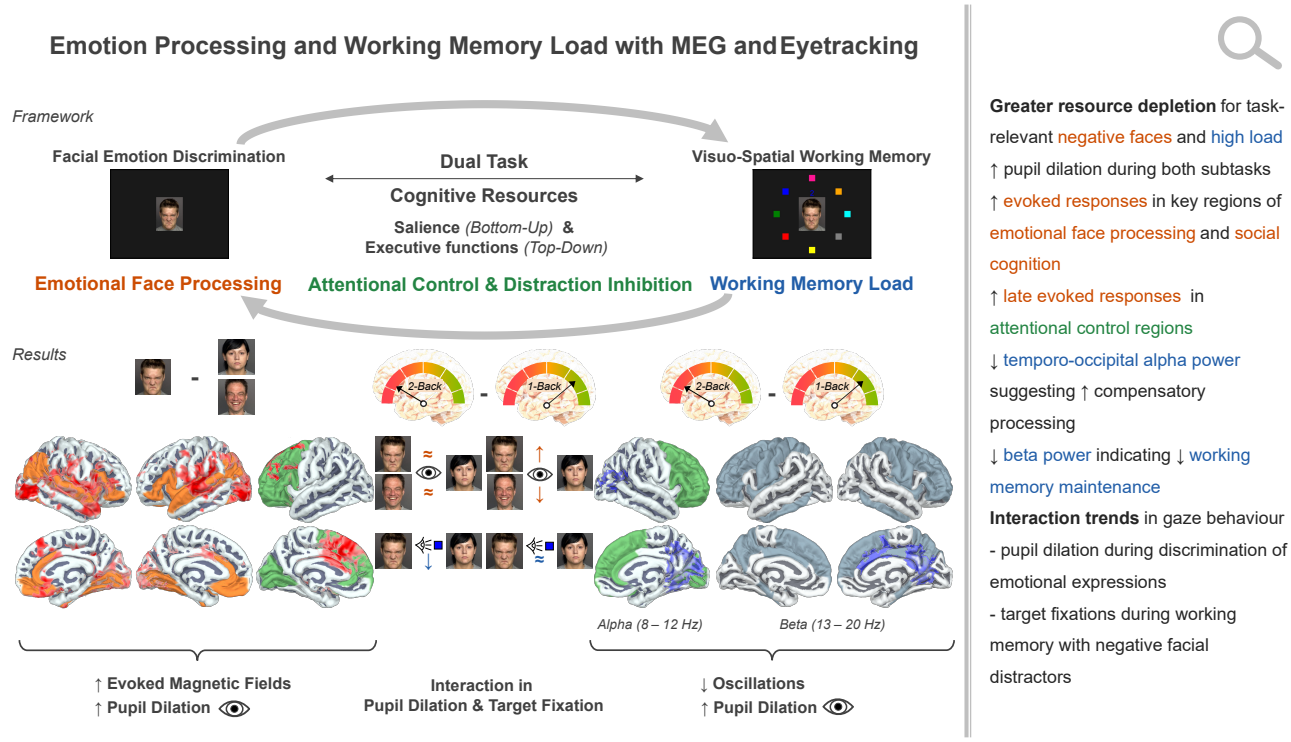
keywords: emotional face processing, working memory load, magnetoencephalography (MEG), eye-tracking, dual task, oscillations, event-related magnetic fields (ERF)

Summary

- Emotionally negative faces enhance evoked brain responses linked to face processing when task-relevant, but not when acting as distractors.
- Working memory load modulates oscillatory signatures only during the encoding and retrieval phase. Decreases in parietal low beta band power suggest diminished working memory maintenance under high load. In contrast, a decline in temporo-occipital alpha power indicates enhanced attention allocation and information processing.
- Neurophysiological and behavioural findings suggest greater resource depletion and reduced attentional control for negative facial expressions and high working memory load.
- Pupil dilation during emotion discrimination and target fixation during the working memory encoding/retrieval phase indicate an interaction between load and emotion.

Figure 2.1

Graphical Abstract



Note. We investigated emotional face processing and working memory in a dual-task MEG-eyetracking study. Task-relevant negative faces increased evoked magnetic fields in regions associated with emotional face processing and attentional control, while high working memory load decreased alpha and beta power in temporo- and parieto-occipital areas, suggesting greater resource depletion under these conditions. Gaze behaviour was affected by an interaction between working memory load and emotional expression.

2.2 Introduction

In the past, cognitive and affective-emotional processes have mostly been studied separately, ignoring their inherent interdependence in human nature (Cromheeke & Mueller, 2014; Pessoa, 2008). In the last decade, this perspective has shifted towards the study of functional brain networks and mechanisms underlying the interaction of emotion and cognition (Brockhoff et al., 2022; Schweizer et al., 2019).

One assumption of this research is that cognitive resources, including attentional control and the capacity to process and retain information, are limited (Baddeley, 1992; Cowan, 2017; Fougnie & Marois, 2006; Wickens, 2014). Attentional control facilitates goal-directed behaviour by guiding perception, processing, and response selection (Mackie et al., 2013; E. K. Miller & Cohen, 2001), while working memory (WM) encodes, maintains, and retrieves task-relevant information (Baddeley, 1992, 1996; Eskikurt et al., 2025; Oberauer, 2019). In dual tasks incorporating WM, a secondary task (e.g., discrimination task) and task-irrelevant stimuli

(e.g., a negative face) can divert attentional and processing resources from the primary task (e.g., maintaining multiple items in WM). This causes the risk of task interference and decay of stored information, especially as subtask demands increase (Cowan, 2005, 2017; Navon & Miller, 2002; Tombu & Jolicoeur, 2003; Watanabe & Funahashi, 2014).

2.2.1 Emotional Face Processing and Categorisation in a Dual Task

Certain stimuli, such as faces, appear to be prioritised in processing regardless of task relevance (see Lavie & de Fockert, 2005; Lavie et al., 2004; but cf. Pessoa et al., 2002; Pessoa et al., 2005; L. Wang et al., 2016). Emotionally negative faces have been proposed to directly and automatically attract attention – even independently of available resources – due to their relevance for adaptive behaviour (Adolphs, 2003; Dolan, 2002; Eimer & Holmes, 2007; Roesch et al., 2010; Schindler & Bublatzky, 2020; Schupp et al., 2006; Vuilleumier, 2005). This bottom-up phenomenon, referred to as emotional attention (Vuilleumier, 2005; Vuilleumier & Huang, 2009), enhances an early event-related neurophysiological component (ERP) named P1/P100 (see Eimer & Holmes, 2007; Schupp et al., 2006; but cf. Schindler & Bublatzky, 2020).

However, further evidence suggests that emotions in expressions are processed more robustly at a later configural stage (see Schindler & Bublatzky, 2020, for review; Hinojosa et al., 2015; Skinner & Benton, 2010), and compete for cognitive resources during ongoing tasks (Hinojosa et al., 2015; Schindler & Bublatzky, 2020). At this configural stage, emotional compared to neutral faces, amplify a component emerging approximately 170 milliseconds after face onset (N/M170; Schindler & Bublatzky, 2020). The M170 is localised in the fusiform face area (FFA; Deffke et al., 2007; Halgren et al., 2000; Kanwisher et al., 1997; Monroe et al., 2013; Pizzagalli et al., 2002, see Figure 2.2A). Its deflection is more pronounced in the right hemisphere (Bentin & Deouell, 2000), and correlates with reaction times in emotion discrimination tasks (Monroe et al., 2013).

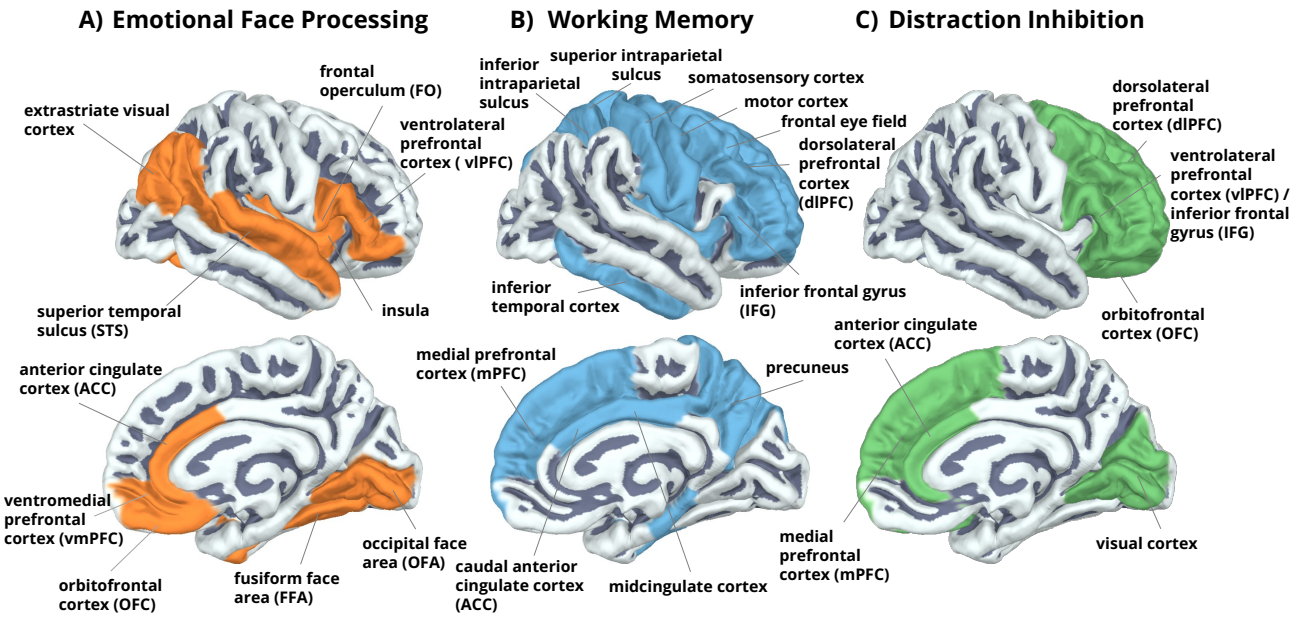
The theory of prototype-referenced shape encoding was proposed to explain mechanisms underlying the processing of configural facial features (Leopold et al., 2001, 2006). In contrast to bottom-up mechanisms (Vuilleumier, 2005; Vuilleumier & Huang, 2009), this theory suggests a top-down guided contrastive mechanism that compares figural variations with an internal face template. Deviations from this template enable the discrimination of facial configurations, including facial expression (Skinner & Benton, 2010). Hence, the difficulty of discriminating a face increases with its resemblance to the internal template (top-down effect). In cases of high similarity, amplified processing and greater attention towards the face may contribute to more accurate differentiation.

Figure 2.2A illustrates key brain areas involved in emotional face processing (Posamentier & Abdi, 2003; Vuilleumier & Pourtois, 2007; Vuilleumier et al., 2003; Weidner et al., 2024; Winston et al., 2003). Some of these regions are linked to the mirror neuron system, and facilitate the interpretation of expressions (Carr et al., 2003; Haxby et al., 2000; Montgomery & Haxby, 2008; Pitcher et al., 2008; Said et al., 2011).

Studies have demonstrated that increased perceptual or WM load can attenuate event-related components associated with emotional face processing (F. Cao et al., 2022; van Dillen & Derks, 2012; L. Wang et al., 2016; P. Yang et al., 2015; but cf. Gläscher et al., 2007; Müller-Bardorff et al., 2018; see Brockhoff et al., 2022, for review), as well as the ability to discriminate facial expressions (F. Cao et al., 2022; Lim et al., 2014).

Figure 2.2

Functional Brain Networks Involved in a Dual Task Alternating Between Facial Emotion Discrimination and Visuo-Spatial Working Memory Load



Note. Key regions from which activity can be measured with magnetoencephalography (MEG) are coloured in orange, blue, and green for the subprocesses of (A) emotional face processing, (B) working memory, and (C) distraction inhibition.

Moreover, van Dillen and Koole (2009) reported a behavioural interaction between facial emotion and WM load, with longer reaction times for angry compared to happy faces only during low but not high load (see also van Dillen & Derks, 2012). The authors attributed this phenomenon to increased attentional demands caused by the negative information. This interference effect is mitigated during high load due to increased attentional control towards the WM task to maintain task performance.

MacNamara et al. (2012) appears to be the only study to examine neurophysiological evidence alongside gaze behaviour. Their study employed fearful and neutral facial distractors during a WM task with two load levels. The authors observed more fixations for neutral compared to fearful faces and high compared to low WM load, but no interaction effect between WM load and emotional distraction on gaze-related measures. The faces were always presented as distractors during the WM maintenance phase. This leaves the question open how WM load and emotionally negative and positive expressions affect gaze behaviour in scenarios where faces are (1) context-specific distractors or targets, and (2) presented not only during maintenance but also simultaneously during WM encoding and retrieval. Fixation duration and pupil size seem to reliably increase with emotional arousal (see Skaramagkas et al., 2021, for review), and provide useful measures for investigating interference, as indicated by increased fixation duration and pupil size for negative facial distractors and targets across WM phases. Unfortunately, MacNamara et al. (2012) did not include these gaze

measures in their analysis.

2.2.2 Working Memory Load in a Dual Task

WM requires an interplay of attention, perception, and response selection (see Figure 2.2B; see D'Esposito & Postle, 2015; Kim, 2019; Rottschy et al., 2012, for meta-analyses).

Neuroimaging studies reported increased activation in the anterior insula, medial prefrontal cortex (mPFC) and inferior frontal gyrus (IFG) with increasing load (Cairo et al., 2004; Rottschy et al., 2012; Tomasi et al., 2007, see Figure 2.2B). However, during high WM load, where capacity reaches its limits, neural activation can be attenuated in the primary motor and sensory cortices, precuneus, and anterior temporal lobe (X. Chen et al., 2023; Michels et al., 2010, see Figure 2.2B). Such reduced activation during high compared to low WM load appears to be correlated with performance and an attenuated representation of task-relevant information (Leung et al., 2004; Michels et al., 2010). Previous magneto- and electrophysiological research suggests that WM load modulates oscillatory power in different frequency bands (see Pavlov & Kotchoubey, 2022, for review): As WM load increases, increased frontal theta and decreased parietal alpha and beta power were found (Scharinger et al., 2015, 2017). However, the role and direction of oscillatory modulations in the alpha and beta band and their link to increasing visual WM load remain inconclusive (Y. Chen & Huang, 2015; Pavlov & Kotchoubey, 2022; Proskovec, Heinrichs-Graham, & Wilson, 2019; Proskovec, Wiesman, et al., 2019). Notably, Pavlov and Kotchoubey (2022) reported that in over 80% of the reviewed studies, WM load modulated alpha- and beta-band oscillations in the same direction, either both increasing or decreasing. In their MEG study, Proskovec, Heinrichs-Graham, and Wilson (2019) found that high WM load enhanced occipital alpha activity during maintenance but decreased it during encoding. Notably, high-load WM simultaneously caused greater alpha suppression in parietal cortices during maintenance, contrasting with load-related alpha power increases in occipital areas (see also Michels et al., 2010). The authors concluded that whether alpha band power increases or decreases with increasing load depends on the specific brain region (occipital vs. parietal) and WM phase (encoding vs. maintenance; Proskovec, Heinrichs-Graham, & Wilson, 2019). Alpha oscillations are assumed to act as a gating mechanism for top-down allocation of attention to task-relevant and suppression of task-irrelevant neural activity (Bonnefond & Jensen, 2012; Boonstra et al., 2013; Fries et al., 2001; Jokisch & Jensen, 2007; Klimesch et al., 2007; Proskovec, Wiesman, et al., 2019; Schroeder et al., 2018). The neural suppression protects task-relevant information and activity from interference (Y. Chen & Huang, 2015; Gevins et al., 1997; Pavlov & Kotchoubey, 2022). In line with this assumption, temporo-occipital alpha band modulations appear to correlate with individual performance differences (Bonnefond & Jensen, 2012; Proskovec, Wiesman, et al., 2019).

At the behavioural level, WM retrieval accuracy declines and reaction time increases as WM load increases (MacNamara et al., 2012; Tavares et al., 2016; van Dillen & Derks, 2012; L. Wang et al., 2016; P. Yang et al., 2015). Increasing load also affects gaze behaviour, leading to increased fixation durations and a larger pupil size (Skaramagkas et al., 2021).

Emotional Distractor Inhibition During Working Memory

Several studies have shown that prefrontal areas play a role in goal-directed attention control and resource allocation to suppress effects of distracting task-irrelevant emotional information (Dolcos & Denkova, 2014; García-Pacios, Garcés, et al., 2015; García-Pacios et al., 2017; Iordan et al., 2013; Ochsner & Gross, 2005; Yoon et al., 2006). Compared to neutral distractors, emotional distractors appear to attenuate and even disrupt WM-related neural processes (see Schweizer et al., 2019, for a meta-analysis; and Dolcos & Denkova, 2014; Dolcos et al., 2011, for reviews). Regulatory frontal areas that counteract these distraction effects comprise dorsal areas, including the dorsolateral PFC (dlPFC), mPFC, anterior cingulate cortex (ACC), as well as ventral regions, including the ventrolateral PFC (vlPFC), orbitofrontal cortex (OFC), and IFG (Ochsner & Gross, 2005, see Figure 2.2C).

Studies investigating oscillatory processes reported an increase in posterior alpha power in response to visual distractions, suggesting enhanced top-down suppression of irrelevant neural activity (Bonnefond & Jensen, 2012; Boonstra et al., 2013; Jokisch & Jensen, 2007; Klimesch et al., 2007; Schroeder et al., 2018).

A magnetoencephalography (MEG) study has yielded insights into the temporal dynamics of emotional image distraction inhibition in a delayed-recognition paradigm (García-Pacios, Garcés, et al., 2015; García-Pacios et al., 2017). These temporal dynamics revealed that functional connectivity of frontal and parietal regions was reduced in early time windows due to the detection of emotional distractors (50 – 150 ms after distractor onset). This was followed by the involvement of prefrontal control processes to suppress distractions and re-establish the fronto-parietal WM processes (250 – 460 ms after distractor onset; García-Pacios, Garcés, et al. 2015; García-Pacios et al. 2017). The pattern was particularly pronounced for emotionally negative distractions (García-Pacios, Garcés, et al., 2015; García-Pacios et al., 2017). Behaviourally, García-Pacios, Del Río, et al. (2015) reported increased reaction times and reduced accuracy during negative compared to positive and neutral distraction, whereas no differences were observed between positive and neutral images. As one of the few studies assessing subjectively perceived valence and arousal of distractor images, they found that participants rated positive images highest in valence, followed by neutral and negative images. In contrast, arousal ratings were highest for negative images, followed by positive and neutral images (García-Pacios, Del Río, et al., 2015). Other behavioural studies have reported interaction effects between WM load and the valence of distractors. WM performance declined during high but not low load scenarios for negative compared to positive and neutral distractors (Li et al., 2012; Tavares et al., 2016, for image distractors; cf. Schweizer et al., 2019, for a meta-analysis; and cf. van Dillen & Derks, 2012; MacNamara et al., 2012; L. Wang et al., 2016; P. Yang et al., 2015, for facial distractors).

2.2.3 Research Question

Taken together, existing evidence on interaction effects between WM load and emotional expressions primarily stems from electrophysiological (Brockhoff et al., 2022; Schindler & Bublatzky, 2020; van Dillen & Derks, 2012) or neuroimaging studies (Gläscher et al., 2007; Müller-Bardorff et al., 2018), often without the inclusion of gaze behaviour (but see MacNamara et al., 2012) and with either poor spatial or temporal

resolution. Most of the studies did not examine interaction effects on a face-related task (but see van Dillen & Derks, 2012; van Dillen & Koole, 2009), and none used a task explicitly emphasising the emotional expressions. Some studies employed only one WM load level (García-Pacios, Del Río, et al., 2015; García-Pacios, Garcés, et al., 2015; García-Pacios et al., 2017). Others were limited to two emotion conditions (negative vs. positive in van Dillen & Derks, 2012; negative vs. neutral in MacNamara et al., 2012; Tavares et al., 2016) or used images as emotional stimuli (García-Pacios et al., 2017; Li et al., 2012; Tavares et al., 2016). In addition, all existing studies focused on distraction effects during WM maintenance (García-Pacios et al., 2017; Li et al., 2012; MacNamara et al., 2012; Tavares et al., 2016; van Dillen & Derks, 2012; P. Yang et al., 2015) without investigating effects during WM encoding/retrieval. Finally, electrophysiological studies on emotional face processing under different load levels have reported inconsistent findings regarding which components and processes are modulated by an interaction (Brockhoff et al., 2022; Müller-Bardorff et al., 2018; Schindler & Bublatzky, 2020; van Dillen & Derks, 2012; P. Yang et al., 2015).

In sum, the evidence for neurophysiological, gaze-related, and behavioural effects of multiple WM load levels and phases, as well as facial distractors including negative, neutral, and positive expressions, remains inconclusive. To address this, the present study examines the effects of WM load and emotion, and specifically their interaction, on (a) neurophysiological dynamics, (b) oscillatory signatures, (c) gaze behaviour, (d) performance, and (e) subjective ratings during facial emotion discrimination and WM in a dual task. For the WM load manipulation, we employed an n-back paradigm with two load levels: 1-back (low WM load; LW) and 2-back (high WM load; HW). The n-back task is widely used to investigate WM (Kim, 2019; Schmiedek et al., 2014) and is well-suited for dual-task paradigms (e.g., Kimura & Matsuura, 2023; Unni et al., 2017). Our experimental paradigm required participants to maintain n-back-related information in WM while simultaneously performing the emotion discrimination subtask (i.e., facial emotion discrimination phase with WM maintenance). To assess facial emotion effects, happy (high valence; HV), neutral (average valence; NV), and angry (low valence; LV) expressions were used as target stimuli in the discrimination task. In the subsequent visuo-spatial WM encoding/retrieval phase, the same faces appeared as distractors while participants compared and updated n-back target information (see Figure 2.3). Brain activity and gaze behaviour were simultaneously recorded using whole-head MEG and eye-tracking.

Given that previous research has only investigated isolated aspects of interacting emotion processing and WM load in the visual domain (regarding stimulus material, number of load levels, and time point of distraction), the current study aims to provide a more comprehensive understanding through the dual-task paradigm. It integrates multimodal evidence to investigate the circumstances under which interaction effects occur and how they manifest in emotional face processing and visuo-spatial WM.

Table 2.1

Summary of the Hypotheses, Organised by Measure (A – E), Task Phase (Facial Emotion Discrimination With WM Maintenance and Visuo-Spatial Working Memory With Encoding/Retrieval), and Effect Type (Main and Interaction Effect)

H.ID	a) MEG ERF	b) MEG Oscillation	c) Gaze Behaviour	d) Performance	e) Ratings
Facial Emotion Discrimination Phase					
	Main Effect of Emotional Expressions (LV - NV and HV - NV)				
1	modulated HV[1] modulated LV[1]	x	↓ fixation duration HV[2,3] ↑ fixation duration LV[2,3]	↓ errors HV[3] ≃ errors LV[3]	↑ valence HV[4] ↓ valence LV[4]
2	↑ early processing (∼ P1) HV[1] ↑ early processing (∼ P1) LV[1]		↓ fixation count HV[2,3] ↑ fixation count LV[2, 3]	↓ reaction time HV[3] ≃ reaction time LV[3]	↑ arousal HV[4] ↑ arousal LV[4]
3	↑ configural processing (∼ M170) HV[1] ↑ configural processing (∼ M170) LV[1]		↓ pupil dilation HV[2] ↑ pupil dilation LV[2]		↓ effort HV[*e] ↑ effort LV[*e]
4	↑ evaluative processing (∼ LPP) HV[1] ↑ evaluative processing (∼ LPP) LV[1]				
	Dual-Task Main Effect of WM Load during Maintenance (HW - LW)				
5	x	modulated[5]	↑ fixation duration[2]	↑ errors[5]	↓ valence[*e]
6		↑ frontal θ[5]	↑ fixation count[2]	↑ reaction time[5]	↑ arousal[*e]
7		↓ parietal α/β[5]	↑ pupil dilation[2]		↑ effort[5]
8		↑ occipital α[5]			
	Interaction Effect (LV - NV / HV - NV HW - LW) ⇒ ≃ NV in HW**		**Interaction Effect (HV - NV HW - LW) ⇒ ≃ NV in HW**		
9	x modulated HV[1*e] x modulated LV[1]	modulated[4,5,7]	↑ fixation duration[6,*e]	↑ errors[6]	↓ valence[*e]
10	↓ early processing (∼ P1) HV[1*e] ↓ early processing (∼ P1) LV[1]	↓ frontal θ[4,5]	↑ fixation count[6,*e]	↑ reaction time[6]	↑ effort[*e]
11	↓ configural processing (∼ M170) HV[1*e] ↓ configural processing (∼ M170) LV[1]	↑ parietal α/β[4,5]	↑ pupil dilation[6,*e]		
12	↓ evaluative processing (∼ LPP) HV[1*e] ↓ evaluative processing (∼ LPP) LV[1]				
Visuo-spatial WM Encoding/Retrieval Phase					
	Main Effect of WM Load during Encoding/Retrieval (HW - LW)				
13	x	modulated[5]	↑ target fixation[2]	↑ errors[5]	↑ effort[5]
14		↑ frontal θ[5]	↑ target onset[2]	↑ reaction time[5]	
15		↓ parietal α/β[5]	↑ pupil dilation[2]		
16		↑ occipital α[5]	↓ face fixation[*e]		
	Depleted Cognitive Resources and Attentional Control ⇒ Interaction Effect HW - LW only for LV***				
17	modulated[4,7]	modulated[4,7]	↓ target fixation[7]	↑ errors[7]	↑ effort[7]
18	↑ early processing (∼ P1)[1,7]	↓ frontal θ[4,5,7]	↑ face fixation[7]	↑ reaction time[7]	
19	↑ configural processing (∼ M170)[1,7]	↑ parietal α/β*[4,5,7]			
20	↑ evaluative processing (∼ LPP)[1,7]	↓ occipital α*[4,5,7]			

Note. H.ID: Hypothesis ID; Main hypotheses are highlighted with blue-coloured, bold row numbers, and blue-coloured sub-headings. ERFs: event-related fields; LV: low valence; NV: neutral valence; HV: high valence; HW: high WM load, LW: low WM load. LPP: late positive potential. *e Hypotheses are marked as exploratory due to a lack of existing research; ** Effects are assumed to be stronger for LV compared to HV [4,7]; *** The combination of LV and HW is assumed to enhance gaze-related, behavioural, and subjective WM load correlates without changing the effect's direction. [1] Brockhoff et al., 2022; Schindler & Bublatzky, 2020; [2] Skaramagkas et al., 2021; [3] Becker & Rheem, 2020; Eastwood et al., 2003; Weidner et al., 2024; [4] García-Pacios, Del Río, et al., 2015; [5] Bonnefond & Jensen, 2012; Michels et al., 2010; Pavlov & Kotchoubey, 2022; Proskovec, Heinrichs-Graham & Wilson, 2019; Proskovec, Wiesman, et al., 2019; Scharinger et al., 2015; Scharinger et al., 2017; [6] van Dillen & Derks, 2012; van Dillen & Koole, 2009; [7] Schweizer et al., 2019. Hypotheses highlighted in green are supported by our results.

Hypotheses

Table 2.1 provides an overview of the hypotheses investigated in this study, structured by measure, dual-task phase, and effect type. In subsequent sections, hypotheses are referenced by their respective IDs.

Based on previous research, we hypothesised that the processing of emotional facial expressions would modulate event-related magnetic fields (ERFs; H.1–4a and H.17–20a), while WM load would manifest in changes of oscillations (during maintenance; H.5–8b; and encoding/retrieval; H.13–16b).

Facial Emotion Discrimination Phase. During the discrimination, we expected a main effect of emotional expression on ERFs associated with emotional face processing (H.1–4a). Positive faces should lead to a processing advantage in gaze-related measures (H.1–4c), performance (H.1–2d) and ratings (H.1–3e). Concurrent high WM load induced by the n-back subtask is assumed to modulate oscillatory power in frequency bands linked to WM maintenance (H.5–8b) and affect gaze-related measures (H.5–8c), performance (H.5–8d), and ratings (H.5–8e). ***Interaction:*** An interaction effect with WM load is hypothesised to manifest in the form of attenuated emotion-related ERF modulations (H.9–12a) and maintenance-related oscillatory signatures (H.9–12b). This effect should be particularly visible for negative faces (García-Pacios, Garcés, et al., 2015; García-Pacios et al., 2017; Schweizer et al., 2019). For positive faces, a high WM load should mitigate any processing advantages observed in gaze-related measures and performance (H.9–11c–e).

Visuo-spatial WM Encoding/Retrieval Phase. During the n-back phase, we predicted a main effect of WM load on oscillatory signatures (H.13b), characterised by increased frontal theta power (H.14b) and decreased parietal alpha and beta power (H.15b). To regulate effects of task-irrelevant facial distractors, posterior (occipital and temporal) alpha power should suppress irrelevant neural activity associated with emotional face processing, assuming cognitive resources for attentional control are available (H.13a,16b). Gaze-related and behavioural measures should reflect goal-directed attention allocation (H.13–14,16c), and an increase in cognitive effort during high WM (H.13–16c–e). ***Interaction:*** However, for negative facial expressions combined with a high WM load, we anticipated an interaction effect with task interference due to heightened processing demands and depleted resources (H.17–20a–e). Consequently, due to loss of attentional control, we expected to observe enhanced ERFs related to task-irrelevant emotional face processing during the n-back encoding/retrieval phase (H.17–20a,c), along with reduced WM-related frontal theta (H.18b) and increased parietal alpha and beta power (H.19b), as well as decreased regulatory posterior alpha power (H.20b).

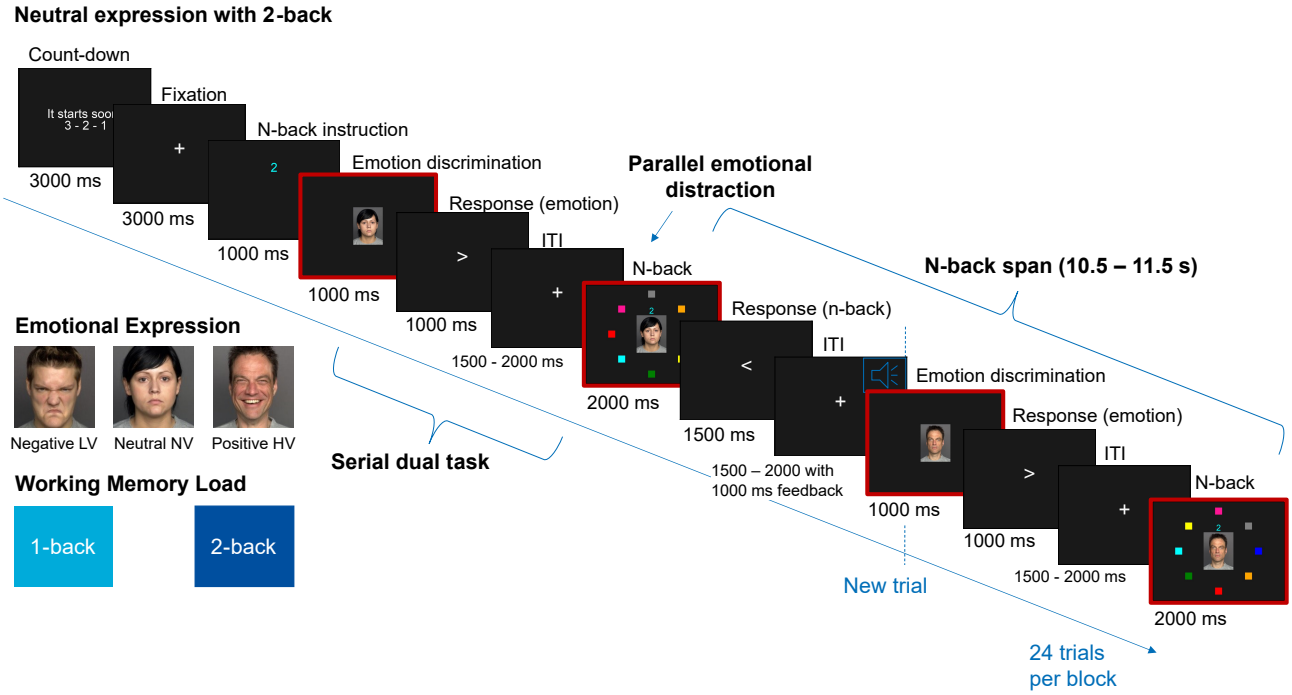
2.3 Methods and Materials

2.3.1 Participants

47 healthy volunteers (mean age $= 24.87$ years, $SD = 3.28$, range: 19 to 33 years, 27 females, 20 males) participated in the study. A power analysis was conducted, indicating a required sample size of at least 36 participants to detect an effect of small to medium size ($\eta_p^2 = 0.03$) with an α of 0.05, power $(1 - \beta)$

Figure 2.3

Dual-Task Experiment Alternating Between Facial Emotion Discrimination With Three Emotional Expressions (LV: Low Valence, NV: Neutral Valence, and HV: High Valence) and a Colour-Based Spatial N-Back Subtask With Two Working Memory (WM) Load Levels (1- and 2-Back)



Note. ITI: Inter-stimulus interval. WM maintenance is required during the emotion discrimination phase. WM encoding of the new target position and retrieval (i.e., matching of the target position with n trials back) co-occur during the n-back phase. Across a block, emotional expression and WM load level remain the same. Facial stimuli are task-relevant in the emotion discrimination phase, while they are task-irrelevant in the n-back phase.

of 0.8, and estimated correlation of repeated measures at $r = 0.5$.

All participants were right-handed, had normal or corrected-to-normal vision, and reported no colour blindness, neurological diseases, psychiatric disorders, or consumption of psychoactive medication or drugs. They gave written informed consent before the experiment and received monetary compensation for their participation. The experiment was approved by the Commission for Research Impact Assessment and Ethics at the Carl von Ossietzky University, Oldenburg, Germany (Ref: EK/2018/070) and conducted in compliance with the Declaration of Helsinki.

For the MEG source-space analyses, we excluded participants with technical issues ($n = 5$), missing anatomical T1 due to claustrophobic attacks in the MRI scanner ($n = 3$), and a very low n-back performance of below $50\,\%$ accuracy ($n = 1$). This exclusion resulted in 38 MEG datasets (mean age $= 24.87$ years, $SD = 3.36$, age range: 19 to 33 years, with 22 females and 16 males). For the eye-tracking analysis, seven participants were excluded due to insufficient calibration or data quality ($N = 40$, mean age $= 24.68$ years,

$SD = 3.07$, age range: 19 to 33 years, with 25 females and 15 males).

2.3.2 Procedure and Material

The experiment comprised 18 pseudo-randomised blocks, divided into three rounds, with six blocks per round and 24 trials within each block (see Supplementary Figure 6.1). Within each round, the six conditions (two WM levels x three emotional expressions) were presented in random order. Each block started with a countdown followed by a fixation cross presented for 3,000 ms. Afterwards, participants had to perform a dual task with two subtasks in each trial. First, they were given a facial emotion discrimination task followed by a visuo-spatial n-back task (Figure 2.3). At the end of each block, participants rated their overall perceived effort using an adapted version of the NASA Task Load Index (NASA-TLX) subscale (Hart & Staveland, 1988), as well as the overall arousal and valence using the Self-Assessment Manikin (SAM) subscales (Bradley & Lang, 1994). Participants underwent two practice blocks at the beginning of the experiment to ensure sufficient task comprehension.

2.3.3 Facial Emotion Discrimination Phase

Face images used for emotion discrimination were obtained from the validated FACES database (Ebner et al., 2010) after permission was granted. The database comprises 513 naturalistic Caucasian faces. Each identity was depicted in three images with either a happy (high valence, HV or positive), angry (low valence, LV or negative), or neutral expression (neutral/average valence, NV; see Figure 2.3). Images were balanced for each condition regarding gender and age group (young, middle-aged, and elderly faces). Within each trial, a face image was presented for 1,000 ms in the middle of the screen, and participants had to discriminate whether the face had a neutral or emotional expression. The facial expression condition (LV, NV, or HV) was constant throughout a block, but identities were randomly selected from the pool of 171 faces without repetition of the same image within the block.

2.3.4 Visuo-Spatial N-Back Phase

The colour-based visuo-spatial n-back task was an adapted version of von Lühmann et al. (2019). Eight coloured squares (red, magenta, blue, turquoise, green, yellow, orange, and grey) were arranged equidistant in a circle. In the middle of the circle, the static face image of the previous facial emotion discrimination phase was repeated to achieve a simultaneous emotion-based interaction during the n-back encoding/retrieval phase. Above the face image, a coloured target number indicated the current target colour and n-back level (Figure 2.3). The coloured target number remained the same over a whole block and was initially presented at the beginning of each block for 1,000 ms (Figure 2.3; n-back instruction in Supplementary Figure 6.1). The colour of each square changed trial-wise within the block. For each trial, they had to compare whether the position of their target colour was either the same as n trials before (same n-back position) or not (different n-back position). Participants performed two WM load levels – a low WM load condition with a 1-back and a high one with a 2-back. Since the n-back interval of remembering the position between trials was quite long, with 10.5 to 11.5 s,

we decided against choosing a high WM load condition above 2-back.

2.3.5 Response Format

We used two response pads with the buttons positioned below the index fingers of the left and right hand. To avoid a fixed button-to-response mapping, resulting in an anticipation effect during the face image and n-back presentation, mapping changed on each trial and was indicated by the direction of an arrow (Figure 2.3). The arrow was presented for 1,000 ms after the emotion discrimination phase and 1,500 ms after the n-back phase. It pointed to the button mapped to an emotional facial expression or the same n-back position. In case of an emotional facial expression or the same n-back position, participants had to press the button the arrow pointed to as quickly as possible. In the cases of a neutral face or different n-back position, they had to press the button opposite the arrow point. The response interval was followed by an inter-stimulus interval (ITI) with a jittered length of 1,500 to 2,000 ms. In the 1,000 ms of the ITI after the n-back encoding/retrieval phase, they received auditory feedback via in-ear headphones on whether their n-back response was correct or not. We did not provide feedback on emotion discrimination to avoid disrupting subsequent emotion-based effects.

2.3.6 Data Acquisition

The experimental dual task was created and presented using Python 3.7 and PyGame (v. 1.9.6). We used a parallel port and Expyriment (v. 0.10.0) to send synchronisation and response triggers to the MEG system. The dual task was displayed using rear projection on a screen measuring 750 x 428 mm inside the MEG chamber with a resolution of 1400 x 1050 pixels (Panasonic PT-DS 12 KE, 60 Hz refresh rate). The back-projection screen was positioned 115 cm away from the participant's eyes. The screen covered $15.07°$ of visual angle in width and $8.73°$ in height from the centre. The face stimuli were centrally presented as rectangles on a dark grey background (#191919), each extending $3.30°$ of visual angle in height and $3.46°$ in width. The circle of squares had a radius of $7.60°$ (outer) and $6.41°$ (inner) of visual angle in height, and $9.95°$ (outer) and $8.40°$ (inner) of visual angle in width from the centre.

Eye-Tracking

We recorded gaze behaviour during the experiment using a MEG-compatible infrared eye-tracking device (EyeLink 1000 Plus; SR Research Ltd., Ottawa, Canada) with a 1 kHz sampling rate. A 9-point calibration was performed at the beginning of each experiment, and a drift correction at the beginning of each block. The average calibration error was kept below $0.5°$ of visual angle ($max < 1.0°$).

Magnetoencephalography

Neuromagnetic signals were recorded using a 306-channel whole-head MEG system (Elekta Neuromag Triux, Elekta Oy, Helsinki, Finland) with 102 magnetometer and 204 orthogonal planar gradiometer sensors. The MEG system was located inside a magnetically shielded chamber (Vacuumschmelze, Hanau, Germany). The dewar was positioned at $68°$ with participants seated underneath the MEG sensors in an upright position.

Five head position indicator coils (HPI) were attached to the participant's head to enable continuous position tracking during the recording. We digitised coil positions along with anatomical landmarks (nasion, left; LPA; and right pre-auricular; RPA; points) and at least 200 head-shape samples (Whalen et al., 2008) using a Polhemus Fastrak (Polhemus, Colchester, VT, USA) to later co-register the MEG data with the structural T1 MRI scans. The MEG signals were recorded without internal active shielding, at a sampling rate of 1 kHz, and with an analogue online band-pass filtering between 0.1 Hz and 330 Hz.

Magnetic Resonance Imaging (MRI) Acquisition

To analyse the MEG data in source space, two structural T1-weighted MRI scans were obtained from each participant, which were averaged to get a better signal-to-noise ratio (SNR). Images were acquired using a Siemens Magnetom Prisma 3.0 Tesla whole-body MRI machine (Siemens, Erlangen, Germany) with a 3D T1-weighted sequence (MPRAGE, TR = 2000 ms, TE = 2.07 ms, flip angle = $9°$, voxel size = 0.75 x 0.75 x 0.75 mm^3, GRAPPA = 2, field of view = 240 x 240, 224 sagittal slices, fat-saturated, TA = 7:45 min). The T1 images were segmented into specific brain tissues (i.e., white matter, brain, skull, and skin), and individual brain surfaces were reconstructed for source localisation using FreeSurfer (v. 6.0.0; Dale et al., 1999; Fischl et al., 1999).

2.3.7 MEG Analysis

The analyses were performed using Python 3.9 and MNE-Python (v. 1.3; Gramfort et al., 2013).

Preprocessing in Sensor Space

To mitigate the influence of external noise, we applied a Maxwell filter to the MEG data, as implemented in MNE-Python (Taulu & Simola, 2006; Taulu et al., 2005) with the default settings (Lin = 8, Lout = 3, correlation limit between inner and outer subspaces = 0.98; Taulu & Simola, 2006; Taulu et al., 2005). The filter uses spatiotemporal signal space separation (tSSS) to decompose the neuromagnetic signal into spatiotemporal components originating from inside and outside the sensor helmet. Magnetic interference that does not originate from brain sources can thereby be attenuated. The Maxwell filter further allows to interpolate bad channels, transform the data into a common coordinate system with equal origin for all rounds, and compensate for head movements by transforming the signals to the initial head position based on the continuous HPI coil tracking. For three subjects, the continuous HPI coil tracking could not be properly recorded due to a technical issue. In these cases, no head movement correction was applied. The head movement correction quality was overall high, with a goodness of fit for the head position indicator coils exceeding 0.99 on a scale of 0 to 1 across all participants. The average 3D head movement distance to the initial point, calculated from the indicator coils, was 3.04 mm ($SD = 1.78$ mm), and $47.09\ \%$ of the distances were below 2 mm. Additionally, the average 3D head movement distance between consecutive samples (sampled at 1 kHz) was 0.18 mm ($SD = 0.13$ mm), with $97.85\ \%$ of the distances falling within a 2 mm threshold. After applying the Maxwell filter, raw data were downsampled to 100 Hz (Fourier method with a boxcar window; as implemented in scipy v. 1.11.4) and band-pass filtered using a 4^{th}-order infinite impulse response (IIR) Butterworth filter. In the

time-domain analyses, a narrow frequency band was chosen with cut-off frequencies set at 0.1 Hz and 20 Hz. In the frequency domain analyses, the chosen cut-offs were 0.1 Hz and 42 Hz.

We analysed two time windows of interest: the time interval starting from the presentation of the face during emotion discrimination, and from the presentation of the n-back wheel with concurrent facial distractor during the visuo-spatial WM encoding/retrieval (Figure 2.3; red highlighted frames).

The filtered signals were segmented into epochs, each lasting from 200 ms before stimulus onset to 1,000 ms after the onset. For analysis of oscillatory power during the n-back encoding/retrieval phase, a larger time interval was selected, extending from 200 ms before to 2,000 ms after the stimulus onset. We concatenated epochs from the three rounds and exclusively selected the magnetometer data for further analyses. We focus our analysis on magnetometers because they allow to probe deeper sources (Hämäläinen et al., 1993), such as orbitofrontal regions, the insula, or the cingulate cortex. These areas were revealed as key regions in prior studies investigating cognitive control mechanisms and emotional processing (Dong et al., 2024; García-Pacios et al., 2017; Iordan et al., 2013).

To remove cardiac-, blink-, and muscle-related artefacts, we performed a semi-automated independent component analysis (ICA) on the epoched data using the FASTER pipeline (v. 1.2; Nolan et al., 2010) and MNE-Python (Gramfort et al., 2013). We performed the ICA with components capturing at least $99\ \%$ of the explained variance ($M = 55.24$, $SD = 3.41$ components for the time-based analysis and $M = 57.57$, $SD = 2.88$ for the frequency-based analysis) and using the extended infomax algorithm (T.-W. Lee et al., 1999). Components contaminated with eye movements were manually selected after visual inspection of the topography, time course, and power spectrum (Chaumon et al., 2015; Hipp & Siegel, 2013). Afterwards, the automated FASTER algorithm was applied to remove further contaminated components with features suggested by Nolan et al. (2010). Electrocardiac signals were reconstructed from the MEG sensors to remove cardiac-related artefacts via the FASTER algorithm. For the time-based analysis, an average of 6.62 components ($SD = 1.17$; $min = 4$, $max = 9$) were removed per participant, while for the frequency-based analysis, an average of 7.26 components ($SD = 1.38$; $min = 5$, $max = 11$) were removed before back-projecting the signals into sensor space.

In the last step, we performed an epoch-wise baseline correction by subtracting the mean of the 200 ms window before the stimulus onset. We excluded the initial epoch from each 1-back block and the first two epochs from each 2-back block. This was necessary because the n-back load had not fully developed during these trials, resulting in a lower WM load as intended in the respective condition.

Source Space Transformation

We reconstructed and localised the sources of the neural activity from the MEG sensor data (Hämäläinen et al., 1993).

To model the conductivity in the brain, we chose the numerical method of a boundary element model (BEM; Mosher et al., 1999) with the FreeSurfer watershed tessellation algorithm (Ségonne et al., 2004). A single-shell head model with the triangulation of the inner skull was used to compute the geometry information in the form of the BEM solution with a conductivity value of 0.3 Siemens/m.

The source space was created using a uniformly distributed grid of dipoles with positions and orientations according to the MNI305 (Montreal Neurologic Institute) space (Collins et al., 1994), as implemented in FreeSurfer. We used icosahedron subdivisions with distance-based spacing of 5 mm, yielding approximately 10242 sources per hemisphere and a source spacing of 3.1 mm. This comprises a surface area per source of 9.8 mm^2. A neuroanatomical label was automatically assigned to each grid point based on the Desikan-Killiany atlas (Desikan et al., 2006; Fischl et al., 2004).

We performed the coregistration in a semi-automated approach following Houck and Claus (2020). First, anatomical landmarks in the MRI headspace (i.e., the fiducial points: LPA, RPA, and nasion) were estimated from an MNI305 brain template (fsaverage; Dale et al., 1999; Gramfort et al., 2013) and transformed to each participant's MRI coordinate space. The estimated landmarks were visually inspected for each subject and manually adjusted to enhance accuracy. In the next step, an initial fit, including scaling, translation, and rotation to the MRI head surface, was performed using only the fiducial points (relative weight for nasion = 10, LPA and RPA = 1). Next, head shape points and MRI were automatically aligned using the iterative closest point (ICP) algorithm (6 iterations, nasion weight = 2, other points' weight = 1). In the third step, head shape points with a distance larger than 5 mm to the MRI skin surface were omitted. In the last step, the ICP algorithm was repeated with 20 iterations (nasion weight = 2). In two participants, the recorded nasion landmark was slightly distorted compared to the rest of the head shape points and fiducials. Therefore, we reduced the nasion weight parameter to 1 and increased the weight of the RPA and LPA to 2 for these participants to acquire a better fitting. Our approach yielded an average distance between head shape points and MRI skin surface of $M = 1.62\,(SD = 1.18)$ mm across participants.

Afterwards, we calculated the leadfield matrix using the coregistration model, source space, and BEM solution for each participant. The noise-covariance matrix was estimated from epoched data using a period of 200 ms before the stimulus onset. Ledoit-Wolf shrinkage was applied for regularisation, and the alpha parameter was optimised through a cross-validated search. The rank was inferred from the Maxwell filtering header.

Subject-wise source reconstruction requires a forward model with rank estimation and a depth prior. As with the noise covariance matrix, ranks were inferred from the data. Additionally, the depth prior was estimated from the data using a weighting exponent of 0.8. We chose the Minimum Norm Estimate (MNE) as the inverse solution method, applying a loose orientation constraint (*weight* = 0.2; Lin et al., 2006). Pooling was performed by taking the norm of the loose orientations. In scenarios where reliable a-priori information about source generators is uncertain – as is often the case in complex cognitive tasks – MNE is a valuable linear inverse method for projecting sensor measurements into source space (Hauk, 2004). It estimates a source distribution with minimum current (L2-norm) that best accounts for the measured data (Hämäläinen et al., 1993). In the time-locked ERF analysis, the inverse operator was applied to the evoked data with an SNR of 3; while in the oscillatory power analysis, it was applied to the epoched data with an SNR of 1 (Gramfort et al., 2013; Lin et al., 2006). The SNR influences the regularisation for scaling the noise-covariance matrix.

Inferential Statistics of the Event-Related Magnetic Fields and Oscillatory Power

For the group-level statistical comparisons, the individual estimated source activity or power spectral density of each condition was morphed to the average FreeSurfer brain template fsaverage (Fischl et al., 1999; Gramfort et al., 2013). This morphing transforms the source space of individual subjects into the same source space.

We computed the power spectral density of the inverse solution from epoched data using the multitaper method with discrete prolate spheroidal sequence (DPSS) windows and a bandwidth of 2 Hz. Cut-offs of the frequency bands range from 4 to 7 Hz for theta, 8 to 12 Hz for alpha, 13 to 20 Hz for low beta, 21 to 29 Hz for high beta, and 30 to 42 Hz for gamma (Pavlov & Kotchoubey, 2022).

We performed the group-level statistic for the time-locked ERFs and oscillatory power per frequency band during the two dual-task phases (Figure 2.3) in two steps: We first tested for significant main effects of emotional expression, WM load, or a significant interaction of the two factors in a repeated-measures analysis of variance (rmANOVA) using spatiotemporal (ERFs) or spatial (oscillatory power) one-sided non-parametric permutation-based clustering (Maris & Oostenveld, 2007) with 5,000 permutations. Permutation-based clustering is a mass-univariate statistical approach that enables data-driven statistical testing across spatial (each vertex) and temporal (each time point) dimensions while controlling for the family-wise error (Maris & Oostenveld, 2007). Traditional univariate ERP/ERF statistics, such as an ANOVA on mean or peak amplitudes at specific locations and within pre-defined time intervals, offer higher statistical sensitivity but provide limited temporal and spatial resolution of effects (Groppe et al., 2011; Pernet et al., 2015). A mass-univariate approach is particularly advantageous when component latencies fluctuate due to experimental factors, such as complex stimulus material (Bentin et al., 1999). In permutation-based clustering, a cluster is considered to have a significant effect if the sum of *F*-values surpassed a pre-defined threshold – here the 95thpercentile ($\alpha = .05$) – of the *F*-value distribution in randomised data (Maris & Oostenveld, 2007). To determine which group comparison differed significantly, we performed two-sided post-hoc *t*-tests on the respective comparisons of the significant main or interaction effect with permutation-based clustering as a second step. The analysis window was confined to the significant time window from the rmANOVA effect to reduce the number of samples along the temporal dimension while preserving spatiotemporal insights (Groppe et al., 2011). In case of a significant main effect, we averaged data over the other main effect and estimated the source activity as described above for the post-hoc *t*-statistics. Significant clusters were projected on a 3D brain (fsaverage), visualising the statistical values (i.e., *F*-values and *t*-values) per cluster.

2.3.8 Eye-Tracking and Behavioural Analysis

Gaze Behaviour

We processed the eye-tracking data by interpolating gaps of missing samples in the time series. This was done using a cubic spline method, with a tolerance for maximum data loss of 75 samples. Subsequently, the time series were smoothed with a median-based rolling window of 20 ms.

Pupil dilation was initially recorded using an arbitrary unit, defined as the number of pixels on the

eye-tracking camera that represented the pupil. To convert this measurement into millimetres, we measured an artificial eye with a known, fixed pupil area of 8 mm to obtain a scaling factor

$$\text{scaling factor} \;=\; \frac{8}{\sqrt{2107}} \;=\; 0.174 \tag{2.1}$$

and applied the scaling as follows:

$$\text{Pupil dilation (mm)} \;=\; \text{Scaling factor} \times \sqrt{\text{pupil dilation (AU)}} \tag{2.2}$$

While we partially compensated for drifts in the centred eye position at the start of each block, we observed a downward shift along the y-axis over time. To counteract this bias, we applied a subject-wise corrective offset to all trials. The offset was calculated for each subject by measuring the deviation between the screen centre and the average y-coordinate of fixations within a ring-shaped region of interest (ROI) surrounding all squares during the n-back encoding/retrieval phase. To reduce the influence of outliers, deviations exceeding 100 pixels were truncated, which affected $34.14\ \%$ of cases. The average correction offset across subjects was $M\ =\ 69.97$ ($SD\ =\ 31.52$) pixels, which corresponds to $M\ =\ 1.17°$ ($SD\ =\ 0.53°$) in visual angles. For the analyses, we used fixations detected by the SR Research online parser and extracted the following parameters: Mean fixation duration, the total count of fixation on the face ROI, and averaged pupil dilation during fixations during both dual-task phases, as well as the onset and total count of fixation on the square at the target position and of the target colour during the n-back phase. The ROI of the face stimulus was defined as a centred ellipse with a width radius of $5.93°$ and a height radius of $4.85°$ of visual angle (see Supplementary Figure 6.6). The ROIs for the squares were defined as circles positioned around the square with a radius of 73 pixels (Supplementary Figure 6.6). To account for family-wise error across the ten models, we corrected the α threshold for the statistical models using the Bonferroni method.

Performance and Ratings

Behavioural data comprised performance measures of made errors in percentage (including false responses and misses) and reaction time (measured in seconds) during the trials, as well as the subjective ratings of perceived valence, arousal, and effort provided at the end of each block.

Inferential Statistics of Gaze Behaviour, Performance, and Ratings

We analysed the main effects of emotional expression and WM load, as well as interactions on gaze behaviour, task performance, and ratings using linear mixed models (Baayen et al., 2008; Bates et al., 2015), as implemented in the toolbox pymer4 (v. 0.8.0; Jolly, 2018). Data for each dependent variable were *z*-standardised, and outliers exceeding the 95thpercentile were removed. Additionally, we included the participant variable as random intercepts in the models to account for non-systematic variations among individuals. The Satterthwaite approximation was applied to adjust the degrees of freedom for heteroscedasticity (using the *anova* method in lmer4; v. 1.1-14; Bates et al., 2015). As post-hoc analyses of significant effects, we

performed non-parametric bootstrapping of group comparisons with 5,000 iterations. Grand averages and their 95thpercentile confidence intervals (CI) of the comparisons were plotted as boxplots. Comparable to the MEG analyses, comparisons with their mean's CI not including zero were considered significant (Cumming & Finch, 2005).

2.4 Results

2.4.1 Behavioural Results of Performance and Subjective Ratings

Linear mixed models assessed whether WM load modulates a processing advantage of positive faces and a disadvantage of negative faces in facial emotion discrimination (van Dillen & Derks, 2012; van Dillen & Koole, 2009, H.9–10d–e). For n-back retrieval, they examined whether high WM load and negative faces deplete cognitive resources, increasing WM task interference, compared to neutral and positive faces under high WM load (van Dillen & Derks, 2012; van Dillen & Koole, 2009, H.17–18d–e). Model results are provided in detail in the Supplementary Table 6.1.

Errors and Reaction Times

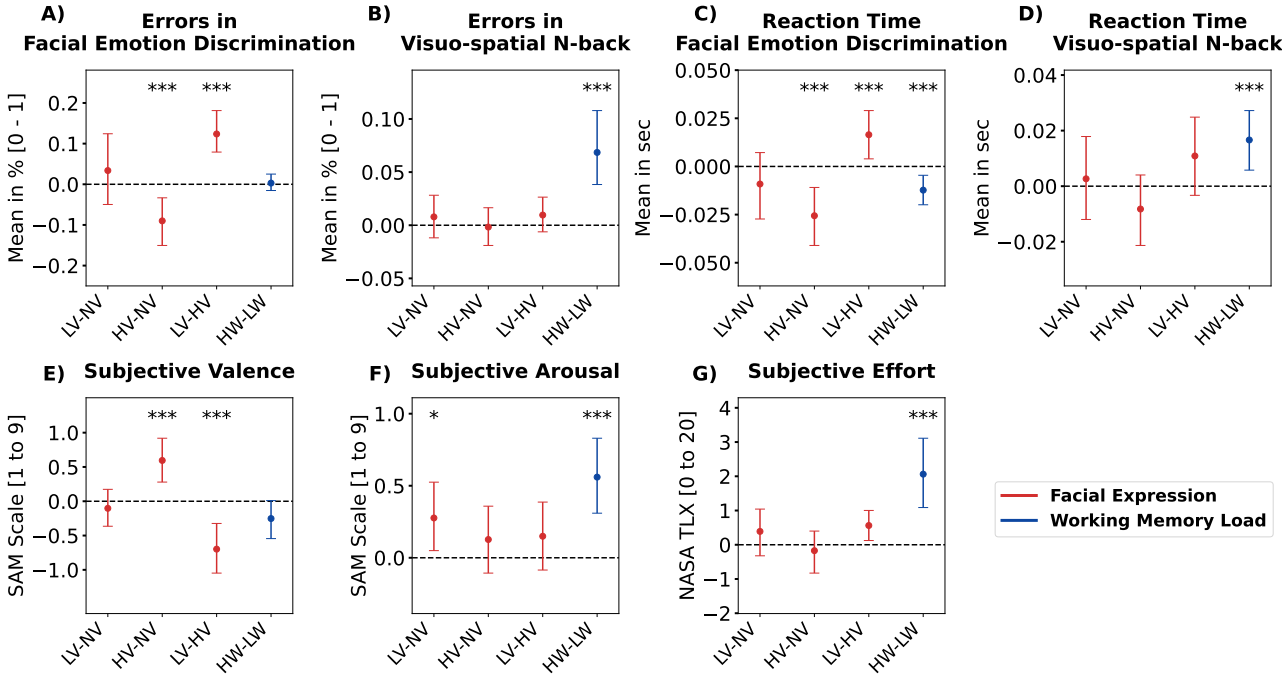
Consistent with the subtask-specific manipulation, we observed a significant main effect of emotional expression ($F_{(2,\ 230)} = 24.52$, $p < .001$) on discrimination errors, and a significant main effect of WM load ($F_{(1,\ 230)} = 68.27$, $p < .001$) on retrieval errors in the n-back subtask. In line with our hypothesis (H.1d), more errors were made during the discrimination of negative and neutral compared to positive faces (LV – HV: $M = 0.124$; $95\ \%\ CI\ [0.079,\ 0.181]$; HV – NV: $M = -0.090$; $95\ \%\ CI\ [-0.151,\ -0.033]$; Figure 2.4A). During the n-back retrieval, high WM load in the 2-back resulted in more errors compared to the 1-back (HW – LW: $M = 0.069$; $95\ \%\ CI\ [0.038,\ 0.108]$; Figure 2.4B; H.13d). There was neither a significant main effect of WM load for emotion discrimination ($F_{(2,\ 230)} = 0.05$, $p = .826$; H.5d), nor a significant interaction in the subtasks (discrimination H.9d: $F_{(2,\ 230)} = 0.30$, $p = .741$; n-back retrieval H.17d: $F_{(2,\ 230)} = 0.01$, $p = .988$).

Likewise, post-hoc comparisons of a significant main effect of emotional expression ($F_{(2,\ 230)} = 15.27$, $p < .001$) revealed longer reaction times when discriminating negative and neutral compared to positive faces (LV – HV: $M = 0.017$; $95\ \%\ CI\ [0.004,\ 0.029]$; HV – NV: $M = -0.026$; $95\ \%\ CI\ [-0.041,\ -0.011]$; Figure 2.4C; H.2d). During the n-back retrieval, we found a significant main effect of WM load ($F_{(1,\ 230)} = 15.21$, $p < .001$), with increased reaction times during high WM load (HW – LW: $M = 0.017$; $95\ \%\ CI\ [0.006,\ 0.027]$; Figure 2.4D; H.14d).

Surprisingly, a significant main effect of WM load ($F_{(1,\ 230)} = 9.91$, $p = .002$) indicated shorter reaction times during emotion discrimination, regardless of facial expression (HW – LW: $M = -0.0123$; $95\ \%\ CI\ [-0.020,\ -0.005]$; Figure 2.4C; H.6d). To examine whether this decline reflects a speed-accuracy trade-off under high WM load, we performed a follow-up analysis. No significant relationship was observed between reaction time changes and accuracy under high WM load (Spearman rank correlation: $r_s = .107$,

Figure 2.4

Post-Hoc Comparisons of the Main Effects of Emotional Expression (Red) and WM Load (Blue) for Behavioural Performance, That Is Errors and Reaction Time, as Well as Subjective Valence, Arousal, and Effort



Note. Coloured dots and error bars represent the bootstrapped grand averages and their Bonferroni-corrected 2.5thand 97.5thconfidence interval (CI) across participants. LV: low valence; NV: neutral valence; HV: high valence; HW: high WM load; LW: low WM load. Significance level from the linear mixed models: *** for $p < .001$, ** for $p < .01$, * for $p < .05$.

$p = .475$). However, irrespective of WM load, reaction times were strongly correlated with discrimination errors ($r_s = .453, p < .001$). Similar to the error results, there was no significant interaction between emotional expression and WM load on reaction time in the emotion discrimination ($F_{(2,\ 230)} = 0.11, p = .898$; H.10d), nor n-back retrieval ($F_{(2,\ 230)} = 0.28, p = .754$; H.18d).

Subjective Valence, Arousal, and Effort

The following linear mixed models analysed the effects of emotion expression (H.1–3e) and WM load (H.5–7,13e), as well as their interaction, on perceived overall valence, arousal, and effort across the block (H.9–10,17e). There were significant main effects of WM load ($F_{(1,\ 230)} = 6.27, p = .013$) and emotional expression ($F_{(2,\ 230)} = 18.80, p < .001$) for valence (Figure 2.4E), significant main effects of WM load ($F_{(1,\ 230)} = 41.08, p < .001$) and emotional expression ($F_{(2,\ 230)} = 3.33, p < .038$) for arousal (Figure 2.4F), and a significant main effect of WM load ($F_{(1,\ 230)} = 75.71, p < .001$) for perceived effort

(Figure 2.4G). There was no significant main effect of emotional expression on effort ratings ($F_{(2,\ 230)} = 1.93$, $p = .148$; Supplementary Table 6.1; H.7e), nor were there significant interaction effects on valence, arousal, or effort ratings (Supplementary Table 6.1; H.9–10e).

The post-hoc analyses revealed higher valence ratings for positive compared to negative (LV – HV: $M = -0.698;\ 95\ \%\ CI\ [-1.048,\ -0.324]$) and neutral expressions (HV – NV: $M = 0.595;\ 95\ \%\ CI\ [0.280,\ 0.919]$; H.1e). There was no significant difference in rated valence in scenarios with negative and neutral faces (H.1e). Valence ratings were lower for high compared to low WM load scenarios, indicating that facial expressions were generally perceived more negatively under high WM (HW – LW: $M = -0.252; 95\ \%\ CI\ [-0.544,\ 0.009]$; n.s.; H.5e). Negative compared to neutral faces and high compared to low WM load led to increased subjective arousal (LV – NV: $M = 0.276; 95\ \%\ CI\ [0.050,\ 0.525]$; H.2e; HW – LW: $M = 0.560; 95\ \%\ CI\ [0.310,\ 0.830]$; H.6e).

Regarding the effort ratings, participants evaluated scenarios with high compared to low WM load to be more effortful (HW – LW: $M = 2.062; 95\ \%\ CI\ [1.087,\ 3.112]$; H.7,13e). Scenarios with negative compared to positive faces were rated as more effortful (LV – HV: $M = 0.562; 95\ \%\ CI\ [0.124,\ 1.002]$; $\sim$ H.3e); however, the main effect of emotional expression was not significant ($F_{(2,\ 230)} = 1.93, p = .148$).

In summary, both performance and subjective ratings were significantly affected by emotional expression and WM load. Increased cognitive demands were observed under high WM load and in scenarios with negative and partially with neutral expressions. This was reflected in increased errors, longer reaction times, reduced valence, and increased arousal and effort.

2.4.2 MEG Results

We examined MEG activation differences between the emotional expression and WM load conditions, as well as their interaction, separately during both dual-task phases (highlighted in Figure 2.3).

Event-Related Magnetic Fields of Emotional Face Processing

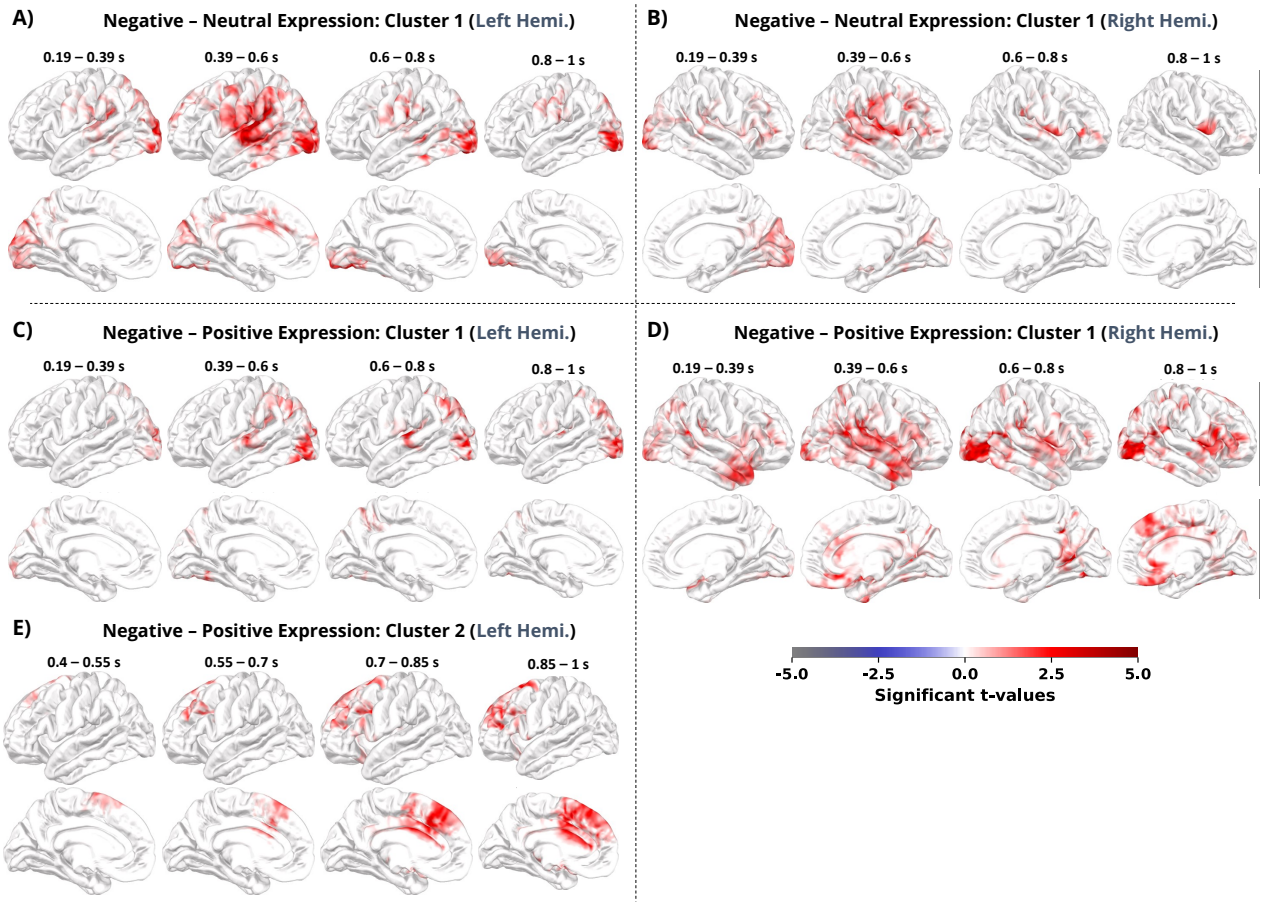
Permutation-based spatiotemporal clustering revealed a significant main effect of emotional expression on ERF amplitudes during facial emotion discrimination ($p_{F-statistic} = .036$) but not during n-back encoding/retrieval, where the face was presented as a distractor for the second time (H.1,13a). We found no effect of WM load on face-related ERFs (H.5a), nor an interaction between WM load and emotional expression (H.9–12a). The significant differences between the ERFs evoked by the different facial expressions started 190 ms after stimulus onset and lasted throughout the analysis time interval (1,000 ms; Supplementary Figure 6.2).

To determine which emotional facial expressions modulate the ERFs and to examine their spatiotemporal distribution, we carried out post-hoc comparisons with a *t*-statistic-based clustering (Figure 2.5). We observed significant clusters with increased ERFs and a similar temporally evolving spatial distribution when contrasting negative and neutral expressions (LV – NV; left hemisphere: $p = .011$; right hemisphere: $p = .006$; Figure 2.5A and B; H.3,4a), and negative and positive facial expressions (LV – HV; left hemisphere:

$p = .011$; right hemisphere: $p < .001$; see Figure 2.5C and D). No ERF difference was found when contrasting positive and neutral facial expressions (H.2–4a).

Figure 2.5

T-Statistic Spatiotemporal Clusters for Significant Comparisons of the Main Effect Emotional Expression. Contrasts Between (A – B) Negative and Neutral Emotional Expressions and (C – E) Negative and Positive Emotional Expressions During Facial Emotion Discrimination



Note. T-values from spatiotemporal clusters were grouped into four equally spaced time intervals and projected onto a 3D brain surface. Lateral (upper) and medial (lower) perspectives are visualised for each cluster and contrast.

In both contrasts, significant bilateral clusters started approximately at 190 ms after the onset of the face and persisted until the end of the presentation. Early significant differences were located at lateral occipital and parietal regions, including the posterior inferior parietal cortex, as well as right fronto-temporal areas, including the anterior temporal lobe and OFC. Around 400 to 600 ms after stimulus onset, effects of pronounced ERF amplitudes for negative facial expressions were located in the lateral occipital cortex and ventral visual stream, temporo-parietal cortex, the anterior insula, posterior superior temporal sulcus (STS),

and anterior temporal lobe (compare to Figure 2.2A). From the medial view (Figure 2.5A and D, second row), we observed slightly increased activation in the caudal anterior and middle cingulate gyrus, superior to the thalamus. These effects were greater in the right hemisphere when comparing negative with positive facial expressions (Figure 2.5D vs. C), and greater in the left hemisphere when comparing negative with neutral expressions (Figure 2.5A vs. B). In the final time interval from 800 to 1000 ms, we observed stronger evoked responses of the lateral occipital cortex and right IFG for negative compared to positive and neutral facial expressions (Figure 2.5A, C, and D).

For the negative – positive expression contrast, we observed a second functional cluster on the left hemisphere, starting at a rather late time point around 400 ms after face onset (Figure 2.5E; $p = .022$). The cluster evolved strongest around the dlPFC, midfrontal gyrus, ACC, and premotor regions after 700 ms from stimulus onset until the end of the presentation.

In summary, our results show enhanced ERFs in the visual cortex and later in fronto-temporal brain regions associated with emotional face processing and social cognition (see also Figure 2.2A) for negative but not positive facial expressions, and only when the emotionally negative faces were task-relevant during the discrimination phase. This modulation was also not affected by the current WM load level (interaction effect; H.9–12a). An exploratory follow-up analysis examining the relationships within and between cluster regions and time intervals can be found in Supplementary Material 6.4.

Oscillatory Power Modulations During the Visuo-Spatial N-Back Encoding/Retrieval Phase

To investigate the effect of WM load on oscillatory signatures (H.5–8,13–16b) and its interactions with emotional distraction (H.9–11,17–20b), spatial clustering was performed for the α, β, and θ frequency bands, as well as, on an exploratory basis, for the γ frequency band during the facial emotion discrimination (i.e., n-back maintenance phase) and n-back encoding/retrieval phase.

We observed significant main effects for WM load in the alpha (8 – 12 Hz; $p_{F-statistic} = .025$; H.15–16b) and low beta (13 – 20 Hz; $p_{F-statistic} = .041$; H.15b) frequency band during the n-back encoding/retrieval phase. No significant spatial clusters were found in the theta, high beta, and gamma frequency bands. Additionally, there was no significant interaction effect elicited by the emotional distractors (H.9–11,17–20b). WM load level modulated oscillatory signatures only event-locked during the encoding/retrieval WM phase (H.13b), but not during the maintenance of the encoded information (i.e., the facial emotion discrimination phase; H.5b).

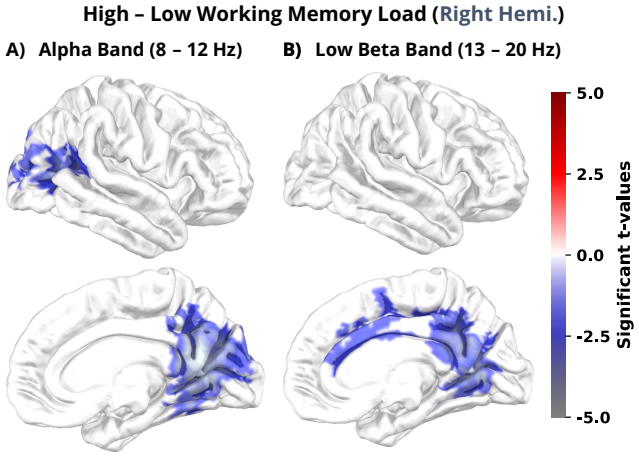
The alpha frequency band cluster exhibited significantly reduced power during high compared to low WM load in the right middle and superior occipital gyrus, lingual gyrus, fusiform gyrus, retrosplenial cortex, and inferior margin of the precuneus ($p_{t-statistic} = .017$; Figure 2.6A).

In the low beta band, we observed decreased oscillatory power during high WM load in the anterior lingual gyrus, inferior margin of the precuneus, inferior parietal lobule, supramarginal gyrus, as well as the anterior and posterior cingulate gyrus ($p_{t-statistic} = .037$; Figure 2.6B, see also Figure 2.2B).

Exploratory follow-up analyses revealed relationships between the alpha band modulations and n-back performance during high WM load. Modulation in the alpha band cluster was positively correlated with retrieval

Figure 2.6

Significant T-Values of the Alpha- and Low Beta-Band Cluster



Note. Significant *t*-values are projected onto a 3D brain surface and visualised from both the lateral (upper row) and medial (lower row) perspectives.

reaction time during high WM load (HW – LW alpha band cluster x HW reaction time: $r_s = .375, p = .020$; see Supplementary Material 6.5, Supplementary Table 6.3, and Supplementary Figure 6.5, for details).

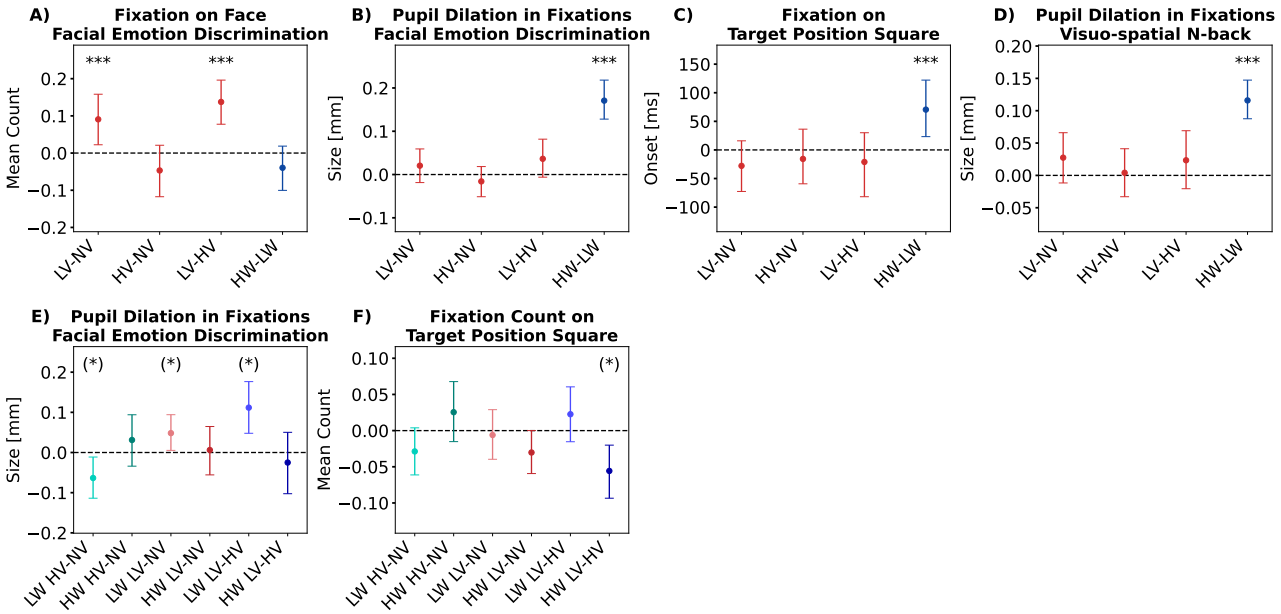
In conclusion, our MEG findings imply that the discrimination of emotional expressions modulated ERFs (H.1,3–4a), while WM load influenced oscillatory brain responses in the alpha and low beta bands during the encoding and retrieval phase (H.13,15–16b). Correlation analysis of oscillatory changes and n-back performance indicates that decreased occipital alpha band power is associated with shorter retrieval reaction times but not with retrieval accuracy, under high WM load. ERFs and oscillatory modulations revealed different processing modes during the discrimination of negative faces, as well as encoding and retrieval of WM under high load. Despite the absence of interaction or interference effects, the observed main effects are likely associated with increased demands on processing resources and attention during negative facial expression discrimination and high-load visuo-spatial WM.

2.4.3 Eye-Tracking Results of the Fixations and Pupil Dilation

We examined interaction effects between emotional facial expressions and WM load on gaze behaviour (i.e., fixation duration and count) and pupil dilation during the facial emotion discrimination (H.5–7c) and n-back encoding/retrieval phase (H.17–18c) using linear mixed models. The statistics for all tested effects are provided in Supplementary Table 6.4, and post-hoc comparisons following a significant effect are illustrated in Figure 2.7.

Figure 2.7

Post-Hoc Comparisons of the Main Effects of Emotional Expression (Red) and WM Load (Blue) of the Gaze-Related Fitted Linear Mixed Models (A – D). Exploratory Comparisons of Interaction Trends are Depicted in (E) and (F)



Note. Coloured dots and error bars represent the bootstrapped grand averages and their Bonferroni-corrected 2.5thand 97.5thconfidence interval (CI) across participants. Visuo-spatial n-back refers to the encoding/retrieval phase. LV: low valence; NV: neutral valence; HV: high valence; HW: high WM load; LW: low WM load. Significance level from the linear mixed models: *** for $p < .001$, ** for $p < .01$, * for $p < .05$, () for significance level without Bonferroni correction of the fitted linear mixed models.

Effects During the Facial Emotion Discrimination Phase

We found a significant main effect of emotional expression on the total count of fixations within the ROI around the face during the emotion discrimination ($F_{(2,\ 195)} = 8.17$, $p < .001$; see Supplementary Figure 6.6, for the display of ROIs in the fixation-related analysis). Negative faces attracted more fixations than both neutral (LV – NV: $M = 0.091$; $95\ \%\ CI\ [0.022,\ 0.158]$; H.2c) and positive faces (LV – HV: $M = 0.138$; $95\ \%\ CI\ [0.078,\ 0.196]$; Figure 2.7A). There was no significant difference in the fixation count between positive and neutral expressions (HV – NV: $M = -0.047$; $95\ \%\ CI\ [-0.117,\ 0.021]$; H.2c), nor were there significant effects of WM load (H.6c), or an interaction with WM load (H.10c) on fixation count. Furthermore, no significant effects were observed for mean fixation duration on the face ROI (H.1,5,9c). In an exploratory follow-up analysis of the main effect of emotional expression on fixation count, we examined which facial regions – the upper part, including the eyes, or the lower part, including the mouth – attracted the most fixations. Results of a linear mixed model with the fixed effects face region (lower vs. upper half) x emotional expression (HV, NV, LV)

revealed a significant difference in fixation counts between face regions ($F_{(1,\ 195)} = 266.423, p < .001$), and mean fixation duration ($F_{(1,\ 195)} = 75.70, p < .001$). More and longer fixations were positioned in the lower face region (lower – upper; count: $M = 1.048; 95\ \%\ CI\ [0.756,\ 1.321]$; duration: $M = 63.473$ ms; $95\ \%\ CI\ [35.994,\ 92.269]$). There was no interaction between facial region and emotional expression (count: $F_{(1,\ 195)} = 1.75, p = .176$; duration: $F_{(1,\ 195)} = 1.05, p = .352$). Analyses including the task phase as a fixed effect in the linear mixed model are provided in Supplementary Material 6.7.1 and 6.7.2.

Sustained Working Memory Load on Pupil Dilation. We observed a sustained WM load main effect on pupil dilation across the dual task ($F_{(1,\ 195)} = 78.99, p < .001$). Pupil dilation increased for high compared to low WM load during the facial emotion discrimination (HW – LW: $M = 0.171$ mm; $95\ \%\ CI\ [0.128,\ 0.219]$; Figure 2.7B; H.7c). There was no significant effect of emotional expression (H.3c), but a strong trend of interaction with WM load (H.11c) during the emotion discrimination ($F_{(2,\ 195)} = 4.46, p < .013$, n.s. after Bonferroni correction on linear mixed model level). Exploratory comparisons showed reduced pupil dilation for positive and increased pupil dilation for negative, compared to neutral faces, only during low WM load (LW HV – NV: $M = -0.063$ mm; $95\ \%\ CI\ [-0.114,\ -0.011]$; LW LV – NV: $M = 0.048$ mm; $95\ \%\ CI\ [0.005,\ 0.094]$; Figure 2.7E).

Effects during the Visuo-Spatial N-Back Encoding/Retrieval Phase

During the n-back phase, we observed a significant main effect of WM load on fixation onset at the target position ($F_{(1,\ 176.25)} = 13.26, p < .001$) and pupil dilation during fixations ($F_{(1,\ 195)} = 16.65$, $p < .001$; Figure 2.7C). During high WM load, the onset of fixation on the target position was delayed (HW – LW: $M = 70.518$ ms; $95\ \%\ CI\ [23.253,\ 122.112]$; H.14c), and pupil dilation was increased (HW – LW: $M = 0.117$ mm; $95\ \%\ CI\ [0.088,\ 0.148]$; Figure 2.7D; H.15c). There was a non-significant trend of interaction between emotional expression and WM load on target fixation count ($F_{(2,\ 195)} = 4.35; p < .014$; H.17c). The exploratory comparisons revealed fewer fixations at the target position during high WM load for negative distractors compared to neutral (HW LV – NV: $M = -0.030; 95\ \%\ CI\ [-0.059,\ 0.000]$; Figure 2.7F) and positive ones (HW LV – HV: $M = -0.056; 95\ \%\ CI\ [-0.093,\ -0.020]$; Figure 2.7F); while no differences were observed between expression conditions during low WM load. There were no significant effects on fixation measures within the face area (H.16,18c).

Taken together, results of the fixation-related correlates and pupil dilation indicated increased cognitive processing demand and required attention recruitment when discriminating negative facial expressions and under high WM load scenarios. Moreover, the WM load level induced by the n-back subtask influenced pupil dilation during facial emotion discrimination, indicating a sustained WM dual-task effect. Sustained WM load interacted with emotional expressions during the discrimination, eliminating differences in pupil dilation between emotional and neutral expressions under high WM load. During the n-back encoding/retrieval phase, an interaction trend suggested that task-oriented gaze behaviour (i.e., fixation on the target position) was reduced for negative emotional distractors, but only under high WM load. Descriptive heatmaps of the fixations during

the emotion discrimination and n-back phase can be found in Supplementary Figures 6.7 and 6.8.

2.5 Discussion

The objective of this whole-head MEG study, combined with eye-tracking, was to address the research gap in understanding the interaction effects of emotional facial expressions and varying WM load levels in a dual-task paradigm. We aimed to identify the circumstances and mechanisms through which emotion and cognitive processing interact during emotion discrimination and visuo-spatial WM. Additionally, we investigated brain mechanisms underlying the task processes and regulatory mechanisms triggered to mitigate indications of task interference when processing and attentional resources are depleted. We analysed MEG source space activation, including ERFs linked to emotional face processing, oscillatory signatures associated with the inhibition of task-irrelevant neural activity and WM load, as well as gaze-related and behavioural responses during both dual-task phases.

As hypothesised, negative facial expressions enhanced spatiotemporal ERFs during the emotion discrimination (Table 2.1; H.1,3–4a). Importantly, our dual-task study suggests that this effect involves top-down mechanisms, as it was only observed when the emotional information was task-relevant. High WM load decreased occipital alpha and parieto-occipital low beta oscillations during the n-back encoding and retrieval phase (Table 2.1; H.13,15–16b). A WM load-based decrease in occipital alpha power was associated with shorter retrieval time, but not increased retrieval accuracy, under high WM load. The relationship between the modulation and performance provides insight into the functional role of alpha band oscillation in enhancing information processing and attention allocation, but not WM maintenance, during increased load. Sustained cross-task effects of high WM load were found on pupil size (H.7c), reaction time (H.6d), and perceived valence (H.5e). Our multimodal study presents converging physiological and behavioural evidence suggesting that the following processes are associated with greater attentional and processing demands: (1) discriminating negative compared to neutral and positive faces, which enhanced evoked brain responses linked to face processing and social cognition, and increased pupil dilation, arousal, and effort; and (2) memory encoding and retrieval during high compared to low load, which decreased posterior alpha and low beta power, and increased pupil dilation, arousal, and effort.

Trends for interactions between WM load and emotion of facial expression were present in both subtasks. During facial emotion discrimination, pupil dilation indicated a processing advantage for positive expressions (decrease) and a disadvantage for negative expressions (increase) during low WM load. However, these emotion-specific processing differences disappeared under high WM load (H.9–11c–e). Moreover, during the n-back encoding/retrieval phase, negative distractors reduced task-related fixations on the target position, but only during high WM load (H.17c). This effect may be associated with reduced attentional control and increased distractibility of negative faces during increased cognitive demands. The interpretation is further supported by the questionnaire result that participants rated blocks with negative facial expressions as more effortful to process (H.3e). By incorporating the neutral expression condition, we were able to extend existing knowledge (van Dillen & Derks, 2012; van Dillen & Koole, 2009), demonstrating that interaction effects are not

solely attributable to interference from negative faces, but also reflect a processing advantage for positive faces under low WM load.

However, we anticipated further interaction effects that our results did not support: During the emotion discrimination, high WM load was hypothesised to mitigate enhanced ERFs of emotional faces and to attenuate oscillatory signatures of WM maintenance (H.9–12a,b). During the n-back encoding/retrieval phase under high WM load, we assumed that angry distractors would interfere with WM due to depleted cognitive resources and reduced attentional control (H.17–20a,b, d,e). This was expected to attenuate the oscillatory signatures of encoding and retrieval (H.17–20b), and elicit enhanced ERFs of emotional face processing (H.17–20a). In the following, we discuss the observed effects in more detail and the context of previous work.

2.5.1 Processing and Discrimination Mechanisms of Emotional Facial Expressions

Enhanced (Socio-)Emotional Evaluation Processes for Negative Facial Expression Discrimination

Negative facial expressions modulated spatiotemporal ERFs with significantly increased activation in fronto-temporal brain networks, the insula-ACC complex, and occipital regions, including the secondary and extrastriate visual cortex. This ERF modulation was only observed during the emotion discrimination and only for negative but not positive faces (H.1,3–4a).

We propose that the observed spatiotemporal ERF clusters may be divided into two functional components. This aligns with the dissociation theory of Ochsner and Gross (2005), who suggests a ventral stream for emotion processing and a dorsal stream for emotion regulation. In our study, the first cluster comprises increased evoked amplitudes in regions associated with emotional face processing (Posamentier & Abdi, 2003; Vuilleumier & Pourtois, 2007; Vuilleumier et al., 2003; Winston et al., 2003, see Figure 2.2A). These enhanced activation patterns of the cluster are observed when contrasting negative to positive but also neutral facial expressions. The engagement of the insula and late reactivation of the occipital cortex suggest that this cluster is also functionally related to more intense processing of negative facial expressions (Carr et al., 2003; Haxby et al., 2000; Montgomery & Haxby, 2008; Pitcher et al., 2008; Said et al., 2011). The intense processing of negative faces is also reflected in increased fixation counts and pupil dilation during the discrimination of negative expressions.

We found a second cluster that appeared only when contrasting negative with positive facial expressions. Since it comprises lateral prefrontal areas and the ACC, this cluster may be linked to goal-directed regulatory mechanisms to overcome potential interference, and disrupt lasting effects of processing and discriminating negative facial expressions (García-Pacios et al., 2017; Ochsner & Gross, 2005). This explanation is supported by recent findings from Dong et al. (2024). In their MEG study, negative image distractors increased activation in the insula during the early phase of the P3b, followed by increased activation in the ACC. They associated this pattern with attentional control to suppress the effect of the distracting negative stimuli.

Although this enhanced ERF linked to attentional control was not modulated by WM load, pupil dilation revealed an interaction between WM load and emotion-specific processing for positive and negative expressions. The effect on pupil dilation aligns with the second ERF cluster and previous findings (van Dillen & Derks, 2012;

van Dillen & Koole, 2009), suggesting processing advantages for positive faces and disadvantages for negative faces under low cognitive demand. However, under high WM load, emotion-specific processing in pupil dilation disappeared (van Dillen & Derks, 2012; van Dillen & Koole, 2009). This finding suggests that the threshold for detecting an interaction between emotion and WM load may be higher at the neural level than in gaze-related measures, and points out the importance of integrating physiological and behavioural measures.

Discrimination Advantage for Positive Facial Expressions

Based on the second ERF cluster and interaction trend in pupil dilation, it can be concluded that negative face processing was enhanced and linked to increased cognitive demands, while positive faces exhibited a processing advantage, at least under low WM load. This is consistent with previous research (Becker et al., 2011; Bucher & Voss, 2019; Calvo & Beltrán, 2013; van Dillen & Koole, 2009; Weidner et al., 2024). In comparing only negative and positive facial expressions, van Dillen and Koole (2009) attributed this difference solely to a negativity bias towards angry faces. However, by including a neutral reference, we could show performance advantages – in the form of reduced errors and reaction time – during the discrimination of positive compared to neutral and negative, but not negative compared to neutral faces. In the ongoing debate on whether there is a processing advantage for negative or positive expressions in visual attention, Xu et al. (2021) concluded in their review that tasks explicitly focusing on emotion processing of photorealistic faces tend to show an advantage for positive expressions. Horstmann et al. (2012) offered a stimulus-level explanation for the processing advantage. The authors demonstrated that the visibility of teeth in photorealistic emotional faces facilitates the search in the face-in-a-crowd paradigm. Given that teeth are visible only in positive facial expressions in the FACES database, participants in our study could have exploited this information for efficient discrimination when distinguishing positive from negative and neutral faces. In line with this, Jehna et al. (2011) reported that neutral faces were more likely to be misjudged as negative than positive faces (see also Albohn et al., 2019; E. Lee et al., 2008; Weidner et al., 2024). Our findings of increased fixations on the lower face ROI support the conclusion that participants relied on information from the mouth region for discrimination. A supplementary fixation analysis, including task phase as a factor, further suggests that this strategy was specific to emotion discrimination and did not occur in the n-back retrieval/encoding phase (see Supplementary Figure 6.7). Interestingly, the potential exploitation of low-level facial characteristics appeared to have a minimal impact on the neural processing in our study: We did not observe modulated spatiotemporal ERFs when contrasting neutral and positive facial expressions, despite their difference in low-level features.

Template-Based Face Encoding as Emotion Discrimination Strategy

Our behavioural and eye-tracking results, thus, support the extension of the prototype-referenced shape encoding theory (Leopold et al., 2001, 2006) to facial expression discrimination (Skinner & Benton, 2010). Hence, behavioural and gaze-related differences between positive and negative facial expressions arise from their differing configural resemblance to the internal template, which is neutral faces (Skinner & Benton, 2010). While positive faces triggered early and efficient discrimination due to strong configural differences to the template, negative faces might have exhibited a higher similarity to neutral faces, resulting in ambiguity

during discrimination (see also the discussion of Weidner et al., 2024). Hence, they required a more thorough examination, which involved recruiting additional neural processes, until clear discriminative features were identified. This would place a greater toll on cognitive resources. The significantly increased number of fixations, and ACC and lateral PFC activation further indicate enhanced attentional engagement with negative faces and the need for control mechanisms to disengage after discrimination (Becker & Rheem, 2020; Eastwood et al., 2003; Horstmann et al., 2006). To summarise, from a template-based facial encoding perspective, the greatest difference in encoding and discrimination fluency can be expected for negative compared to positive facial expressions, at least during low WM load. The latter is associated with efficient and resource-saving discrimination, whereas negative facial expressions appear to be resource-demanding.

The interaction trend in pupil dilation further supports the involvement of top-down attentional mechanisms to discriminate facial expressions. Unlike automatic processes, such deliberate mechanisms can be influenced by the current WM load and available cognitive resources (Schindler & Bublatzky, 2020). It appears that explicit instruction towards emotion classification can further influence whether processing is deliberate or automatic – particularly for positive facial expressions (Rellecke et al., 2012; Xu et al., 2021). This aligns with the absence of any early enhanced ERFs by emotional expressions in our study. Contrary to previous literature (Eimer & Holmes, 2007; Schupp et al., 2006), we did not observe enhanced ERFs for positive facial expressions. One explanation is provided by Almeida et al. (2016), who demonstrated a positive relationship between the modulation of face-related M/EEG components and perceived arousal. Since participants in our study rated angry – but not happy – faces more arousing than neutral ones, the amplified ERFs may be at least partially driven by arousal. In addition to the heightened arousal elicited by angry faces, the increased evolutionary relevance of negative facial expressions, along with their communicative function in signalling potential threat and initiating adaptive responses in social contexts, may also account for the observed processing differences (Adolphs, 2003).

2.5.2 Effects of Cognitive Capacity Limits During High Working Memory Load

Another key finding of our study was a significant main effect of WM load, with high load decreasing oscillatory alpha and low beta band power in posterior regions (H.13,15–16b), and increasing pupil dilation (H.15c), fixation speed (H.14c), the number of errors (H.13d), and reaction times (H.14d) during the encoding and retrieval phase of the visuo-spatial WM n-back subtask. Furthermore, WM load level affected the overall effort, arousal, and even valence experienced across the block (H.5–7,13e). A decline in occipital alpha band power and its WM load-based modulation was associated with faster retrieval reaction times, but not retrieval accuracy, under high WM load (Supplementary Analysis 6.5).

The WM load effects on performance are in line with previous studies (e.g., Leung et al., 2004; Schroeder et al., 2018) and can be explained by insufficient capacity and limited resources to process and maintain task-relevant information.

At the brain level, increased WM load reduced posterior alpha and low beta band power in inferior parietal regions, which strongly overlap with the WM brain network and load-modulated areas (e.g., Lamm et al., 2001; Michels et al., 2010; D. G. V. Mitchell et al., 2007; see also Figure 2.2B). In our study, frontal theta

power did not increase with increasing WM load (cf. Costers et al., 2020; Jensen & Tesche, 2002, H.6, 14b). A possible explanation is that frontal theta signals primarily originate from superficial radial dipole layers, which are reliably detected by EEG but less effectively captured by MEG (Srinivasan et al., 2006).

Modulations in alpha band oscillations have been proposed to guide information processing (Jokisch & Jensen, 2007; Klimesch et al., 2007; Schroeder et al., 2018) with increased alpha power indicating top-down suppression of neural activity, and decreased alpha power indicating cortical engagement. Hence, our results suggest increased visual processing and involvement of the temporo-occipital regions during high compared to low WM load scenarios. An explanation for this phenomenon is that visual processing is enhanced to accommodate the heightened cognitive demand of the more challenging task. This notion is further supported by our correlation results (Supplementary Analysis 6.5), which demonstrate that greater decreases in occipital alpha power are associated with shorter retrieval times under high WM load. Contrary to our hypothesis (H.16b), we conclude that the modulation of occipital alpha band power likely reflects a compensatory mechanism that enhances cortical engagement, facilitating attentional allocation and information processing under high WM load (Jokisch & Jensen, 2007; Klimesch et al., 2007; Schroeder et al., 2018).

The interpretation of load effects on beta oscillatory power is still debated in the literature (Pavlov & Kotchoubey, 2022). Previous studies have proposed a link between increased beta oscillations and WM maintenance (e.g., Y. Chen & Huang, 2015; Deiber et al., 2007; Engel & Fries, 2010). Consistent with this suggestion, we observed a decrease in low beta oscillatory power in key regions associated with WM and spatial cognition (Figure 2.2B), including the inferior parietal areas, retrosplenial cortex, and inferior precuneus (A. S. Mitchell et al., 2018; Vann et al., 2009). We propose that modulatory changes in parieto-occipital low beta band power are related to WM maintenance and the decay of task-relevant information during high load (Y. Chen & Huang, 2015; Deiber et al., 2007; Engel & Fries, 2010; Kopell et al., 2011; Salazar et al., 2012).

To conclude, there is a partial spatial overlap of modulations in the inferior parietal areas between the alpha and low beta band clusters. However, the alpha band cluster includes more temporo-occipital regions; while the low beta band cluster contains almost exclusively inferior parietal regions, which reflect posterior key areas associated with WM (Figure 2.2B). The slightly different spatial patterns and the relationship between alpha band modulations and retrieval performance indicate distinct functional roles of band-specific oscillations in the visuo-spatial WM encoding and retrieval. Decreased occipital alpha power likely represents a compensatory mechanism to enhance information processing under high WM load (Jokisch & Jensen, 2007; Klimesch et al., 2007; Schroeder et al., 2018), whereas reduced parietal low beta band power may reflect diminished WM maintenance and the decay of encoded information (Y. Chen & Huang, 2015; Deiber et al., 2007; Engel & Fries, 2010; Kopell et al., 2011; Salazar et al., 2012).

Sustained Working Memory Load Effects due to Increased Cognitive Demand

Contrary to our initial hypotheses, we did not observe oscillatory modulations locked to the onset of the facial stimuli during emotion discrimination (H.5–8b) nor interactions with facial expressions during both subtasks (H.9–11,17–20b). However, we observed different processing modes reflecting a higher demand on cognitive resources: Negative facial expressions recruited additional neural regions associated with social

cognition and categorisation, as well as amplified activation in the emotional face brain network. High WM load led to decreased WM maintenance and compensatory engagement of occipital regions to enhance attention allocation and information processing.

Due to a greater demand on processing resources, we observed a sustained WM load effect across subtasks: Pupil dilation increased for positive and decreased for negative faces, while reaction time decreased during emotion discrimination with concurrent high compared to low WM load. This pattern suggests that the emotional expression effect on gaze behaviour is reduced during high WM load (see van Dillen & Koole, 2009, for similar behavioural findings). The faster reaction times during high compared to low WM load pointed to quick decision-making in emotion discrimination, presumably driven by enhanced processing resource allocation for the upcoming encoding and retrieval of information from WM.

2.5.3 How Robust are Interaction Effects?

Although some studies have reported interactions between emotion and WM load on neurophysiological and behavioural signatures (e.g., García-Pacios, Garcés, et al., 2015; MacNamara et al., 2012; van Dillen & Derks, 2012), we could not entirely replicate these effects in this dual-task study (H.9–12,17–20a,b). Given the inconsistency in detecting interaction effects in the literature and relatively small reported effect sizes (Brockhoff et al., 2022; Schweizer et al., 2019), some interaction effects may not be robust to variations in experimental design or stimulus material.

A novel aspect less examined in previous studies was the contextual change in task relevance of facial stimuli, specifically their emotional expressions, in our dual task. Contrary to van Dillen and Derks (2012), we aimed to amplify the effect of emotional negative and positive expressions by using a task that explicitly indicated emotion (Rellecke et al., 2012). However, this may have had the opposite effect in the n-back encoding/retrieval phase, reducing distraction and interference during the second presentation of the same, now task-irrelevant face due to prior processing (also discussed by Tavares et al., 2016).

Further, we cannot exclude the possibility that not only emotion effects on ERFs, but also the interaction with WM load is partly influenced by stimulus arousal. This is particularly relevant as most studies did not report subjective ratings (but cf. García-Pacios, Del Río, et al., 2015; García-Pacios, Garcés, et al., 2015). Britton et al. (2006) compared valence and arousal ratings for emotional facial stimuli and images (International Affective Picture System; IAPS; Lang et al., 1997). The IAPS dataset has also been used in studies showing interaction effects for WM and negative stimuli (García-Pacios, Garcés, et al., 2015; Tavares et al., 2016, for behavioural effects). Compared to emotional faces, images were rated significantly more positive for depicted happy and neutral emotions, more negative for negative emotions, and more arousing across all emotion categories (Britton et al., 2006). Based on these findings, faces might be considered weaker distractors due to their lower arousal strength (Britton et al., 2006; Ochsner & Gross, 2005; Tavares et al., 2016). However, Carretié et al. (2012) compared emotional faces and images in their potential to capture attention and observed similar behavioural and neurophysiological responses. Nevertheless, the arousal intensity of the stimuli may have influenced the likelihood of detecting an interaction during discrimination. In addition, arousal levels may reach a ceiling earlier for emotional facial stimuli than for non-facial images (Britton et al., 2006). Future studies

should systematically examine the impact of arousal and stimulus type on interaction effects.

2.5.4 Limitation and Future Directions

The use of the naturalistic FACES database impacted the interpretation of certain effects. We decided against using schematic or morphed emotional faces, which are restricted to an oval shape to remove additional distinctive features, such as hair. However, naturalistic faces carry the risk of introducing confounding low-level features. While these features are negligible when evenly distributed across expressions, they may bias emotion discrimination if they systematically indicate a specific expression (e.g., visible teeth in happy faces). Therefore, our findings of a template-based discrimination strategy should be replicated using schematic faces with carefully controlled low-level features.

As our sample resembled young, highly educated students, the generalisability of the findings to other age groups may be limited.

Concerning source localisation, MNE relies on linear assumptions about the relationship between the electrical activity of the brain and measured signals (Hämäläinen et al., 1993). Interactions of activity from brain regions nearby might not be accurately captured. It also tends to favour sources closer to the sensors when solving the inverse problem (Hämäläinen et al., 1993). By adding a regularisation to capture deeper sources such as insular activation, the estimates obtained sometimes have a spatial spread that is scattered around the actually smaller underlying deep sources.

Our study did not explore individual variations in executive functions, cognitive control over emotionally irrelevant information, or WM capacity. Future research should examine how these individual differences influence the interplay between emotion and cognition in the dual task (e.g., Dolcos & Denkova, 2014; Dolcos et al., 2011). Linking such variations to neural correlates when facing an emotion-cognition interaction could potentially also serve as diagnostic indices from a clinical perspective (Schweizer et al., 2019).

In future investigations, integrating eye-tracking with MEG to perform fixation-related analyses aligned with the onset of fixation on emotional stimuli may enhance sensitivity in detecting time-locked interaction effects (Baccino & Manunta, 2005; Spiering & Dimigen, 2024). Furthermore, additional research could include transcranial magnetic stimulation in the emotion-cognition dual task to increase the likelihood of task interference by inhibiting regulatory mechanisms (Olk et al., 2015).

2.6 Conclusion

In a dual task, we investigated spatiotemporal and oscillatory signatures of emotional face processing and visuo-spatial WM load. Our MEG findings revealed enhanced ERFs that were spatially located across the insula, ACC, and face-specific occipital regions during the discrimination of negative facial expressions but not during the encoding and retrieval of n-back information from WM. Hence, negative faces amplified face processing and social cognition; only when they were task-relevant. Furthermore, when contrasting negative and positive facial expressions, enhanced ERFs were observed in prefrontal regions and the ACC at late time intervals of the discrimination. Both areas are associated with goal-directed executive functions and may be

linked to attentional control, facilitating the dissociation from negative faces to prevent interference with the WM subtask. In addition to these findings, task performance and gaze behaviour demonstrated an advantage in discriminating positive faces, indicated by fewer errors and fixations on the face area, as well as reduced reaction times and subjective effort, and decreased fixations on the face area. An interaction trend in pupil dilation suggested that the processing advantage for positive faces, along with a disadvantage for negative faces, diminished with increasing WM load.

In the n-back encoding/retrieval phase, high WM load attenuated alpha band power in temporo-occipital regions, while low beta band power was reduced in parieto-occipital regions and the cingulate cortex. The spatial localisation of the clusters and their association with retrieval performance suggest that alpha and beta oscillations have different functional roles during increased WM load. Reduced temporo-occipital alpha power indicates compensatory cortical engagement to enhance attention allocation and information processing, while inferior parietal low beta power is proposed to be associated with WM maintenance. At the behavioural level, an interaction trend was observed for negative distractors with reduced fixations on the target position; but only during high WM load. Sustained effects of increased WM load were observed in task performance and gaze behaviour across both subtasks. High WM load reduced the perceived valence of facial expressions and increased pupil size and reaction time during both subtasks.

In conclusion, the study advanced our understanding of when and how negative faces influence event-related magnetic fields, mechanisms underlying emotion discrimination in naturalistic faces, interaction effects in gaze behaviour, and the role of WM-related oscillatory alpha and low beta power under different load levels.

Funding

The research was supported by the Fraunhofer Gesellschaft with the scholarship Fraunhofer TALENTA, the Ministry of Economic Affairs, Labour, and Tourism Baden-Wuerttemberg in the project KI-Fortschrittszentrum Lernende Systeme und Kognitive Robotik, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; RI 1511/3-1 awarded to J. Rieger), and by the Neuroimaging Unit of the Carl von Ossietzky Universität Oldenburg, funded by grants from the DFG (3T MRI INST 184/152-1 FUGG and MEG INST 184/48-1 FUGG).

Declaration of Competing Interests

The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this article. The study was pre-registered on the Open Science Framework (https://osf.io/cg7nm) and the protocol was approved by the ethics committee of the Carl von Ossietzky University, Oldenburg, Germany (Ref. EK/2018/070).

Data and Code Availability

Data supporting the findings of this manuscript are available upon request via the OSF project (https://osf.io/um6vw/). Participants provided informed consent following the Declaration of Helsinki. They were informed that their participation was voluntary, that they could withdraw at any time during the experiment, and that they could request the deletion of their data up until the point of anonymisation. The informed consent also included a section on the publication of anonymised data and aggregated results.

Acknowledgements

We would like to thank Dr. Tina Schmitt for her support during the source reconstruction of the anatomical T1 scans, Dr. Mathias Vukelić for the valuable discussions, Anna Vorreuther for proofreading, and Dr. Andreas Spiegler for the support in the technical set-up and data collection.

References Study I

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, *4*(3), 165–178. https://doi.org/10.1038/nrn1056

Albohn, D. N., Brandenburg, J. C., & Adams, R. B. (2019). Perceiving emotion in the “neutral” face: A powerful mechanism of person perception. In U. Hess & S. Hareli (Eds.), *The social nature of emotion expression* (pp. 25–47). Springer. https://doi.org/10.1007/978-3-030-32968-6_3

Almeida, P. R., Ferreira-Santos, F., Chaves, P. L., Paiva, T. O., Barbosa, F., & Marques-Teixeira, J. (2016). Perceived arousal of facial expressions of emotion modulates the N170, regardless of emotional category: Time domain and time–frequency dynamics. *International Journal of Psychophysiology*, *99*, 48–56. https://doi.org/10.1016/j.ijpsycho.2015.11.017

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baccino, T., & Manunta, Y. (2005). Eye-fixation-related potentials: Insight into parafoveal processing. *Journal of Psychophysiology*, *19*(3), 204–215. https://doi.org/10.1027/0269-8803.19.3.204

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. (1996). The fractionation of working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(24), 13468–13472. https://doi.org/10.1073/pnas.93.24.13468

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: Happy faces, not angry faces, are more efficiently detected in single- and

multiple-target visual search tasks. *Journal of Experimental Psychology: General*, *140*(4), 637–659. https://doi.org/10.1037/a0024060

Becker, D. V., & Rheem, H. (2020). Searching for a face in the crowd: Pitfalls and unexplored possibilities. *Attention, Perception, and Psychophysics*, *82*(2), 626–636. https://doi.org/10.3758/s13414-020-01975-7

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology*, *17*(1), 35–55. https://doi.org/10.1080/026432900380472

Bentin, S., Mouchetant-Rostaing, Y., Giard, M.-H., Echallier, J.-F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*(3), 235–260. https://doi.org/10.1162/089892999563373

Bonnefond, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Current Biology*, *22*(20), 1969–1974. https://doi.org/10.1016/j.cub.2012.08.029

Boonstra, T. W., Powell, T. Y., Mehrkanoon, S., & Breakspear, M. (2013). Effects of mnemonic load on cortical activity during visual working memory: Linking ongoing brain activity with evoked responses. *International Journal of Psychophysiology*, *89*(3), 409–418. https://doi.org/10.1016/j.ijpsycho.2013.04.001

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Britton, J. C., Taylor, S. F., Sudheimer, K. D., & Liberzon, I. (2006). Facial expressions and complex IAPS pictures: Common and differential networks. *NeuroImage*, *31*(2), 906–919. https://doi.org/10.1016/j.neuroimage.2005.12.050

Brockhoff, L., Schindler, S., Bruchmann, M., & Straube, T. (2022). Effects of perceptual and working memory load on brain responses to task-irrelevant stimuli: Review and implications for future research. *Neuroscience & Biobehavioral Reviews*, *135*, 104580. https://doi.org/10.1016/j.neubiorev.2022.104580

Bucher, A., & Voss, A. (2019). Judging the mood of the crowd: Attention is focused on happy faces. *Emotion*, *19*(6), 1044–1059. https://doi.org/10.1037/emo0000507

Cairo, T. A., Liddle, P. F., Woodward, T. S., & Ngan, E. T. C. (2004). The influence of working memory load on phase specific patterns of cortical activity. *Cognitive Brain Research*, *21*(3), 377–387. https://doi.org/10.1016/j.cogbrainres.2004.06.014

Calvo, M. G., & Beltrán, D. (2013). Recognition advantage of happy faces: Tracing the neurocognitive processes. *Neuropsychologia*, *51*(11), 2051–2061. https://doi.org/10.1016/j.neuropsychologia.2013.07.010

Cao, F., Zeng, K., Li, W., Liu, S., Zhang, L., Katembu, S., & Xu, Q. (2022). Influence of scene-based expectation on facial expression perception: The moderating effect of cognitive load. *Biological Psychology*, *168*, 108247. https://doi.org/10.1016/j.biopsycho.2021.108247

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National*

Academy of Sciences of the United States of America, *100*(9), 5497–5502. https://doi.org/10.1073/pnas.0935845100

Carretié, L., Kessel, D., Carboni, A., López-Martín, S., Albert, J., Tapia, M., Mercado, F., Capilla, A., & Hinojosa, J. A. (2012). Exogenous attention to facial vs non-facial emotional visual stimuli. *Social Cognitive and Affective Neuroscience*, *8*(7), 764–773. https://doi.org/10.1093/scan/nss068

Chaumon, M., Bishop, D. V., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*, *250*, 47–63. https://doi.org/10.1016/j.jneumeth.2015.02.025

Chen, X., Sorenson, E., & Hwang, K. (2023). Thalamocortical contributions to working memory processes during the n-back task. *Neurobiology of Learning and Memory*, *197*, 107701. https://doi.org/10.1016/j.nlm.2022.107701

Chen, Y., & Huang, X. (2015). Modulation of alpha and beta oscillations during an n-back task with varying temporal memory load. *Frontiers in Psychology*, *6*, 2031. https://doi.org/10.3389/fpsyg.2015.02031

Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, *18*(2), 192–205. https://doi.org/10.1097/00004728-199403000-00005

Costers, L., van Schependom, J., Laton, J., Baijot, J., Sjøgård, M., Wens, V., De Tiège, X., Goldman, S., D'Haeseleer, M., D'Hooghe, M. B., Woolrich, M., & Nagels, G. (2020). Spatiotemporal and spectral dynamics of multi-item working memory as revealed by the n-back task using MEG. *Human Brain Mapping*, *41*(9), 2431–2446. https://doi.org/10.1002/hbm.24955

Cowan, N. (2005). *Working memory capacity*. Psychology Press. https://doi.org/10.4324/9780203342398

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cromheeke, S., & Mueller, S. C. (2014). Probing emotional influences on cognitive control: An ALE meta-analysis of cognition emotion interactions. *Brain Structure & Function*, *219*(3), 995–1008. https://doi.org/10.1007/s00429-013-0549-z

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *The American Psychologist*, *60*(2), 170–180. https://doi.org/10.1037/0003-066X.60.2.170

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Deffke, I., Sander, T., Heidenreich, J., Sommer, W., Curio, G., Trahms, L., & Lueschow, A. (2007). MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *NeuroImage*, *35*(4), 1495–1501. https://doi.org/10.1016/j.neuroimage.2007.01.034

Deiber, M.-P., Missonnier, P., Bertrand, O., Gold, G., Fazio-Costa, L., Ibañez, V., & Giannakopoulos, P. (2007). Distinction between perceptual and attentional processing in working memory tasks: A study of phase-locked and induced oscillatory brain dynamics. *Journal of Cognitive Neuroscience*, *19*(1), 158–172. https://doi.org/10.1162/jocn.2007.19.1.158

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science*, *298*(5596), 1191–1194. https://doi.org/10.1126/science.1076358

Dolcos, F., & Denkova, E. (2014). Current emotion research in cognitive neuroscience: Linking enhancing and impairing effects of emotion on cognition. *Emotion Review*, *6*(4), 362–375. https://doi.org/10.1177/1754073914536449

Dolcos, F., Iordan, A. D., & Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: A review of evidence from brain imaging investigations. *Journal of Cognitive Psychology*, *23*(6), 669–694. https://doi.org/10.1080/20445911.2011.594433

Dong, X., Cui, L., & Johnson, B. W. (2024). Neural mechanisms for secondary suppression of emotional distractors: Evidence from concurrent electroencephalography–magnetoencephalography data. *Emotion*, *24*(8), 1907–1922. https://doi.org/10.1037/emo0001388

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Perception & Psychophysics*, *65*(3), 352–358. https://doi.org/10.3758/BF03194566

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES – A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*(1), 351–362. https://doi.org/10.3758/BRM.42.1.351

Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia*, *45*(1), 15–31. https://doi.org/10.1016/j.neuropsychologia.2006.04.022

Engel, A. K., & Fries, P. (2010). Beta-band oscillations – Signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165. https://doi.org/10.1016/j.conb.2010.02.015

Eskikurt, G., Duru, A. D., Ermutlu, N., & İşoğlu-Alkaç, Ü. (2025). Evaluation of brain electrical activity of visual working memory with time-frequency analysis. *Clinical EEG and Neuroscience*, *56*(5), 419–431. https://doi.org/10.1177/15500594231224014

Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *8*(4), 272–284. https://doi.org/10.1002/(sici)1097-0193(1999)8:4⟨272::aid-hbm10⟩3.0.co;2-4

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., & Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, *14*(1), 11–22. https://doi.org/10.1093/cercor/bhg087

Fougnie, D., & Marois, R. (2006). Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms. *Psychological Science*, *17*(6), 526–534. https://doi.org/10.1111/j.1467-9280.2006.01739.x

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, *291*(5508), 1560–1563. https://doi.org/10.1126/science.1055465

García-Pacios, J., Del Río, D., Villalobos, D., Ruiz-Vargas, J. M., & Maestú, F. (2015). Emotional interference-based forgetting in short-term memory. Cognitive inhibition of pleasant but not unpleasant biologically relevant distractors. *Frontiers in Psychology*, *6*, 582. https://doi.org/10.3389/fpsyg.2015.00582

García-Pacios, J., Garcés, P., Del Río, D., & Maestú, F. (2015). Early detection and late cognitive control of emotional distraction by the prefrontal cortex. *Scientific Reports*, *5*, 10046. https://doi.org/10.1038/srep10046

García-Pacios, J., Garcés, P., Del Río, D., & Maestú, F. (2017). Tracking the effect of emotional distraction in working memory brain networks: Evidence from an meg study. *Psychophysiology*, *54*(11), 1726–1740. https://doi.org/10.1111/psyp.12912

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385. https://doi.org/10.1093/cercor/7.4.374

Gläscher, J., Rose, M., & Büchel, C. (2007). Independent effects of emotion and working memory load on visual activation in the lateral occipital complex. *The Journal of Neuroscience*, *27*(16), 4366–4373. https://doi.org/10.1523/JNEUROSCI.3310-06.2007

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. S. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267. https://doi.org/10.3389/fnins.2013.00267

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral cortex*, *10*(1), 69–81. https://doi.org/10.1093/cercor/10.1.69

Hämäläinen, M. S., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography – Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*(2), 413–497. https://doi.org/10.1103/RevModPhys.65.413

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183, Vol. 52). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hauk, O. (2004). Keep it simple: A case for using classical minimum norm estimation in the analysis of EEG and MEG data. *NeuroImage*, *21*(4), 1612–1621. https://doi.org/10.1016/j.neuroimage.2003.12.018

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Hinojosa, J. A., Mercado, F., & Carretié, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *55*, 498–509. https://doi.org/10.1016/j.neubiorev.2015.06.002

Hipp, J. F., & Siegel, M. (2013). Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. *Frontiers in Human Neuroscience*, *7*, 338. https://doi.org/10.3389/fnhum.2013.00338

Horstmann, G., Lipp, O. V., & Becker, S. I. (2012). Of toothy grins and angry snarls – Open mouth displays contribute to efficiency gains in search for emotional faces. *Journal of Vision*, *12*(5), 7. https://doi.org/10.1167/12.5.7

Horstmann, G., Scharlau, I., & Ansorge, U. (2006). More efficient rejection of happy than of angry face distractors in visual search. *Psychonomic Bulletin & Review*, *13*(6), 1067–1073. https://doi.org/10.3758/bf03213927

Houck, J. M., & Claus, E. D. (2020). A comparison of automated and manual co-registration for magnetoencephalography. *PLOS ONE*, *15*(4), e0232100. https://doi.org/10.1371/journal.pone.0232100

Iordan, A. D., Dolcos, S., & Dolcos, F. (2013). Neural signatures of the response to emotional distraction: A review of evidence from brain imaging investigations. *Frontiers in Human Neuroscience*, *7*, 200. https://doi.org/10.3389/fnhum.2013.00200

Jehna, M., Neuper, C., Ischebeck, A., Loitfelder, M., Ropele, S., Langkammer, C., Ebner, F., Fuchs, S., Schmidt, R., Fazekas, F., & Enzinger, C. (2011). The functional correlates of face perception and recognition of emotional facial expressions as evidenced by fMRI. *Brain Research*, *1393*, 73–83. https://doi.org/10.1016/j.brainres.2011.04.007

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*(8), 1395–1399. https://doi.org/10.1046/j.1460-9568.2002.01975.x

Jokisch, D., & Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *The Journal of Neuroscience*, *27*(12), 3244–3251. https://doi.org/10.1523/JNEUROSCI.5399-06.2007

Jolly, E. (2018). Pymer4: Connecting R and Python for linear mixed modeling. *Journal of Open Source Software*, *3*(31), 862. https://doi.org/10.21105/joss.00862

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kim, H. (2019). Neural activity during working memory encoding, maintenance, and retrieval: A network-based model and meta-analysis. *Human Brain Mapping*, *40*(17), 4912–4933. https://doi.org/10.1002/hbm.24747

Kimura, T., & Matsuura, R. (2023). The content-dependent effect of the n-back task on dual-task performance. *Behavioural Brain Research*, *452*, 114511. https://doi.org/10.1016/j.bbr.2023.114511

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Kopell, N., Whittington, M. A., & Kramer, M. A. (2011). Neuronal assembly dynamics in the beta1 frequency range permits short-term memory. *Proceedings of the National Academy of Sciences*, *108*(9), 3779–3784. https://doi.org/10.1073/pnas.1019676108

Lamm, C., Windischberger, C., Leodolter, U., Moser, E., & Bauer, H. (2001). Evidence for premotor cortex activity during dynamic visuospatial imagery from single-trial functional magnetic resonance imaging and event-related slow cortical potentials. *NeuroImage*, *14*(2), 268–283. https://doi.org/10.1006/nimg.2001.0850

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings* (tech. rep.) (Retrieved December 10, 2020, from https://acordo.net/acordo/wp-content/uploads/2020/08/instructions.pdf). NIMH Center for the Study of Emotion and Attention, University of Florida. Gainesville, FL.

Lavie, N., & de Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review*, *12*(4), 669–674. https://doi.org/10.3758/BF03196756

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lee, E., in Kang, J., Park, I. H., Kim, J.-J., & An, S. K. (2008). Is a neutral face really evaluated as being emotionally neutral? *Psychiatry Research*, *157*(1-3), 77–85. https://doi.org/10.1016/j.psychres.2007.02.005

Lee, T.-W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, *11*(2), 417–441. https://doi.org/10.1162/089976699300016719

Leopold, D. A., Bondar, I. V., & Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature*, *442*(7102), 572–575. https://doi.org/10.1038/nature04951

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*(1), 89–94. https://doi.org/10.1038/82947

Leung, H.-C., Seelig, D., & Gore, J. C. (2004). The effect of memory load on cortical activity in the spatial working memory circuit. *Cognitive, Affective & Behavioral Neuroscience*, *4*(4), 553–563. https://doi.org/10.3758/cabn.4.4.553

Li, X., Ouyang, Z., & Luo, Y.-j. (2012). The cognitive load affects the interaction pattern of emotion and working memory. *International Journal of Cognitive Informatics and Natural Intelligence*, *6*, 68–81. https://doi.org/10.4018/jcini.2012040104

Lim, S.-L., Bruce, A. S., & Aupperle, R. L. (2014). The influence of a working memory task on affective perception of facial expressions. *PLOS ONE*, *9*(10), e111074. https://doi.org/10.1371/journal.pone.0111074

Lin, F.-H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*, *31*(1), 160–171. https://doi.org/10.1016/j.neuroimage.2005.11.054

Mackie, M.-A., van Dam, N. T., & Fan, J. (2013). Cognitive control and attentional functions. *Brain and Cognition*, *82*(3), 301–312. https://doi.org/10.1016/j.bandc.2013.05.004

MacNamara, A., Schmidt, J., Zelinsky, G. J., & Hajcak, G. (2012). Electrocortical and ocular indices of attention to fearful and neutral faces presented under high and low working memory load. *Biological Psychology*, *91*(3), 349–356. https://doi.org/10.1016/j.biopsycho.2012.08.005

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Michels, L., Bucher, K., Lüchinger, R., Klaver, P., Martin, E., Jeanmonod, D., & Brandeis, D. (2010). Simultaneous EEG-fMRI during a working memory task: Modulations in low and high frequency bands. *PLoS ONE*, *5*(4), e10298. https://doi.org/10.1371/journal.pone.0010298

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Mitchell, A. S., Czajkowski, R., Zhang, N., Jeffery, K., & Nelson, A. J. D. (2018). Retrosplenial cortex and its role in spatial cognition. *Brain and Neuroscience Advances*, *2*. https://doi.org/10.1177/2398212818757098

Mitchell, D. G. V., Nakic, M., Fridberg, D., Kamel, N., Pine, D. S., & Blair, R. J. R. (2007). The impact of processing load on emotion. *NeuroImage*, *34*(3), 1299–1309. https://doi.org/10.1016/j.neuroimage.2006.10.012

Monroe, J. F., Griffin, M., Pinkham, A., Loughead, J., Gur, R. C., Roberts, T. P. L., & Edgar, J. C. (2013). The fusiform response to faces: Explicit versus implicit processing of emotion. *Human Brain Mapping*, *34*(1), 1–11. https://doi.org/10.1002/hbm.21406

Montgomery, K. J., & Haxby, J. V. (2008). Mirror neuron system differentially activated by facial expressions and social hand gestures: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *20*(10), 1866–1877. https://doi.org/10.1162/jocn.2008.20127

Mosher, J. C., Leahy, R. M., & Lewis, P. S. (1999). EEG and MEG: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, *46*(3), 245–259. https://doi.org/10.1109/10.748978

Müller-Bardorff, M., Bruchmann, M., Mothes-Lasch, M., Zwitserlood, P., Schlossmacher, I., Hofmann, D., Miltner, W., & Straube, T. (2018). Early brain responses to affective faces: A simultaneous EEG-fMRI study. *NeuroImage*, *178*, 660–667. https://doi.org/10.1016/j.neuroimage.2018.05.081

Navon, D., & Miller, J. (2002). Queuing or sharing? A critical evaluation of the single-bottleneck notion. *Cognitive Psychology*, *44*(3), 193–251. https://doi.org/10.1006/cogp.2001.0767

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods*, *192*(1), 152–162. https://doi.org/10.1016/j.jneumeth.2010.07.015

Oberauer, K. (2019). Working memory and attention – A conceptual analysis and review. *Journal of Cognition*, *2*(1), 36. https://doi.org/10.5334/joc.58

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–249. https://doi.org/10.1016/j.tics.2005.03.010

Olk, B., Peschke, C., & Hilgetag, C. C. (2015). Attention and control of manual responses in cognitive conflict: Findings from TMS perturbation studies. *Neuropsychologia*, *74*, 7–20. https://doi.org/10.1016/j.neuropsychologia.2015.02.008

Pavlov, Y. G., & Kotchoubey, B. (2022). Oscillatory brain activity and maintenance of verbal and visual working memory: A systematic review. *Psychophysiology*, *59*(5), e13735. https://doi.org/10.1111/psyp.13735

Pernet, C., Latinus, M., Nichols, T., & Rousselet, G. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *Journal of Neuroscience Methods*, *250*, 85–93. https://doi.org/10.1016/j.jneumeth.2014.08.003

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148–158. https://doi.org/10.1038/nrn2317

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neural and emotional stimuli. *Cognitive Brain Research*, *15*(1), 31–45. https://doi.org/10.1016/s0926-6410(02)00214-8

Pessoa, L., Padmala, S., & Morland, T. (2005). Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. *NeuroImage*, *28*(1), 249–255. https://doi.org/10.1016/j.neuroimage.2005.05.048

Pitcher, D., Garrido, L., Walsh, V., & Duchaine, B. C. (2008). Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. *The Journal of Neuroscience*, *28*(36), 8929–8933. https://doi.org/10.1523/JNEUROSCI.1450-08.2008

Pizzagalli, D. A., Lehmann, D., Hendrick, A. M., Regard, M., Pascual-Marqui, R. D., & Davidson, R. J. (2002). Affective judgments of faces modulate early activity (approximately 160 ms) within the fusiform gyri. *NeuroImage*, *16*(3), 663–677. https://doi.org/10.1006/nimg.2002.1126

Posamentier, M. T., & Abdi, H. (2003). Processing faces and facial expressions. *Neuropsychology Review*, *13*(3), 113–143. https://doi.org/10.1023/a:1025519712569

Proskovec, A. L., Heinrichs-Graham, E., & Wilson, T. W. (2019). Load modulates the alpha and beta oscillatory dynamics serving verbal working memory. *NeuroImage*, *184*, 256–265. https://doi.org/10.1016/j.neuroimage.2018.09.022

Proskovec, A. L., Wiesman, A. I., Heinrichs-Graham, E., & Wilson, T. W. (2019). Load effects on spatial working memory performance are linked to distributed alpha and beta oscillations. *Human Brain Mapping*, *40*(12), 3682–3689. https://doi.org/10.1002/hbm.24625

Rellecke, J., Sommer, W., & Schacht, A. (2012). Does processing of emotional facial expressions depend on intention? Time-resolved evidence from event-related brain potentials. *Biological Psychology*, *90*(1), 23–32. https://doi.org/10.1016/j.biopsycho.2012.02.002

Roesch, E. B., Sander, D., Mumenthaler, C., Kerzel, D., & Scherer, K. R. (2010). Psychophysics of emotion: The QUEST for emotional attention. *Journal of Vision*, *10*(3), 4–9. https://doi.org/10.1167/10.3.4

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., Fox, P. T., & Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*, *60*(1), 830–846. https://doi.org/10.1016/j.neuroimage.2011.11.050

Said, C. P., Haxby, J. V., & Todorov, A. (2011). Brain systems for assessing the affective value of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1660–1670. https://doi.org/10.1098/rstb.2010.0351

Salazar, R. F., Dotson, N. M., Bressler, S. L., & Gray, C. M. (2012). Content-specific fronto-parietal synchronization during visual working memory. *Science*, *338*(6110), 1097–1100. https://doi.org/10.1126/science.1224000

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2015). When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology*, *52*(10), 1293–1304. https://doi.org/10.1111/psyp.12500

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2017). Comparison of the working memory load in n-back and working memory span tasks by means of EEG frequency band power and P300 amplitude. *Frontiers in Human Neuroscience*, *11*, 6. https://doi.org/10.3389/fnhum.2017.00006

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex*, *130*, 362–386. https://doi.org/10.1016/j.cortex.2020.06.010

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: Putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, *5*, 1475. https://doi.org/10.3389/fpsyg.2014.01475

Schroeder, S. C. Y., Ball, F., & Busch, N. A. (2018). The role of alpha oscillations in distractor inhibition during memory retention. *European Journal of Neuroscience*, *48*(7), 2516–2526. https://doi.org/10.1111/ejn.13852

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: Event-related brain potential studies. *Progress in Brain Research*, *156*, 31–51. https://doi.org/10.1016/S0079-6123(06)56002-9

Schweizer, S., Satpute, A. B., Atzil, S., Field, A. P., Hitchcock, C., Black, M., Barrett, L. F., & Dalgleish, T. (2019). The impact of affective information on working memory: A pair of meta-analytic reviews of behavioral and neuroimaging evidence. *Psychological Bulletin*, *145*(6), 566–609. https://doi.org/10.1037/bul0000193

Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*(3), 1060–1075. https://doi.org/10.1016/j.neuroimage.2004.03.032

Skaramagkas, V., Giannakakis, G., Ktistakis, E., Manousos, D., Karatzanis, I., Tachos, N. S., Tripoliti, E., Marias, K., Fotiadis, D. I., & Tsiknakis, M. (2021). Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Reviews in Biomedical Engineering*, *16*, 260–277. https://doi.org/10.1109/RBME.2021.3066072

Skinner, A. L., & Benton, C. P. (2010). Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychological Science*, *21*(9), 1248–1253. https://doi.org/10.1177/0956797610380702

Spiering, L., & Dimigen, O. (2024). (Micro)saccade-related potentials during face recognition: A study combining EEG, eye-tracking, and deconvolution modeling. *Attention, Perception, & Psychophysics*, 1–22. https://doi.org/10.3758/s13414-024-02846-1

Srinivasan, R., Winter, W. R., & Nunez, P. L. (2006). Source analysis of EEG oscillations using high-resolution EEG and MEG. *Progress in Brain Research*, *159*, 29–42. https://doi.org/10.1016/S0079-6123(06)59003-X

Taulu, S., & Simola, J. (2006). Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Physics in Medicine and Biology*, *51*(7), 1759–1768. https://doi.org/10.1088/0031-9155/51/7/008

Taulu, S., Simola, J., & Kajola, M. (2005). Applications of the signal space separation method. *IEEE Transactions on Signal Processing*, *53*(9), 3359–3372. https://doi.org/10.1109/TSP.2005.853302

Tavares, T. P., Logie, K., & Mitchell, D. G. (2016). Opposing effects of perceptual versus working memory load on emotional distraction. *Experimental Brain Research*, *234*, 2945–2956. https://doi.org/10.1007/s00221-016-4697-2

Tomasi, D., Chang, L., Caparelli, E. C., & Ernst, T. (2007). Different activation patterns for working memory load and visual attention load. *Brain Research*, *1132*(1), 158–165. https://doi.org/10.1016/j.brainres.2006.11.030

Tombu, M., & Jolicoeur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 3–18. https://doi.org/10.1037//0096-1523.29.1.3

Unni, A., Ihme, K., Jipp, M., & Rieger, J. W. (2017). Assessing the driver's current level of working memory load with high density functional near-infrared spectroscopy: A realistic driving simulator study. *Frontiers in human neuroscience*, *11*, 167. https://doi.org/10.3389/fnhum.2017.00167

van Dillen, L. F., & Derks, B. (2012). Working memory load reduces facilitated processing of threatening faces: An ERP study. *Emotion*, *12*(6), 1340–1349. https://doi.org/10.1037/a0028624

van Dillen, L. F., & Koole, S. L. (2009). How automatic is "automatic vigilance"? The role of working memory in attentional interference of negative information. *Cognition and Emotion*, *23*(6), 1106–1117. https://doi.org/10.1080/02699930802338178

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, *10*(11), 792–802. https://doi.org/10.1038/nrn2733

von Lühmann, A., Boukouvalas, Z., Müller, K.-R., & Adalı, T. (2019). A new blind source separation framework for signal analysis and artifact rejection in functional near-infrared spectroscopy. *NeuroImage*, *200*, 72–88. https://doi.org/10.1016/j.neuroimage.2019.06.021

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011

Vuilleumier, P., Armony, J., Dolan, R., et al. (2003). Reciprocal links between emotion and attention. *Human Brain Function*, *2*, 419–444. https://doi.org/10.1016/B978-012264841-0/50022-6

Vuilleumier, P., & Huang, Y.-M. (2009). Emotional attention: Uncovering the mechanisms of affective biases in perception. *Current Directions in Psychological Science*, *18*(3), 148–152. https://doi.org/10.1111/j.1467-8721.2009.01626.x

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*, *45*(1), 174–194. https://doi.org/10.1016/j.neuropsychologia.2006.06.003

Wang, L., Feng, C., Mai, X., Jia, L., Zhu, X., Luo, W., & Luo, Y.-J. (2016). The impact of perceptual load on the non-conscious processing of fearful faces. *PLOS ONE*, *11*(5), e0154914. https://doi.org/10.1371/journal.pone.0154914

Watanabe, K., & Funahashi, S. (2014). Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. *Nature Neuroscience*, *17*(4), 601–611. https://doi.org/10.1038/nn.3667

Weidner, E. M., Moratti, S., Schindler, S., Grewe, P., Bien, C. G., & Kissler, J. (2024). Amygdala and cortical gamma-band responses to emotional faces are modulated by attention to valence. *Psychophysiology*, e14512. https://doi.org/10.1111/psyp.14512

Whalen, C., Maclin, E. L., Fabiani, M., & Gratton, G. (2008). Validation of a method for coregistering scalp recording locations with 3D structural MR images. *Human Brain Mapping*, *29*(11), 1288–1301. https://doi.org/10.1002/hbm.20465

Wickens, C. D. (2014). The structure of attentional resources. In *Attention and Performance VIII* (pp. 239–257). Psychology Press.

Winston, J. S., O'Doherty, J., & Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage*, *20*(1), 84–97. https://doi.org/10.1016/s1053-8119(03)00303-3

Xu, Q., Ye, C., Gu, S., Hu, Z., Lei, Y., Li, X., Huang, L., & Liu, Q. (2021). Negative and positive bias for emotional faces: Evidence from the attention and working memory paradigms. *Neural Plasticity*, 8851066. https://doi.org/10.1155/2021/8851066

Yang, P., Wang, M., Jin, Z., & Li, L. (2015). Visual short-term memory load modulates the early attention and perception of task-irrelevant emotional faces. *Frontiers in Human Neuroscience*, *9*, 490. https://doi.org/10.3389/fnhum.2015.00490

Yoon, J. H., Curtis, C. E., & D'Esposito, M. (2006). Differential effects of distraction during working memory on delay-period activity in the prefrontal cortex and the visual association cortex. *NeuroImage*, *29*(4), 1117–1126. https://doi.org/10.1016/j.neuroimage.2005.08.024

3 Study II – GAUDIE: Development, Validation, and Exploration of a Naturalistic German AUDItory Emotional Database

Katharina Lingelbach[1,2], Mathias Vukelić[2],
and Jochem W. Rieger[1]

[1] Department of Psychology, University of Oldenburg, Oldenburg, Germany
[2] Fraunhofer Institute for Industrial Engineering IAO, Nobelstraße 12, 70569 Stuttgart, Germany

This article has been published as:
Lingelbach, K., Vukelić, M., & Rieger, J. W. (2024). GAUDIE: Development, validation, and exploration of a naturalistic German auditory emotional database. *Behavior Research Methods, 56*, 2049–2063. https://doi.org/10.3758/s13428-023-02135-z

Author contributions:
Katharina Lingelbach: Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualisation, Writing – original draft, Writing – review & editing. **Mathias Vukelić**: Validation, Writing – review & editing. **Jochem W. Rieger**: Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.

The original publication and this chapter are identical in content.
I hereby confirm that Katharina Lingelbach contributed to the aforementioned study as stated above.



Prof. Dr. Jochem W. Rieger
Oldenburg, 30.01.2026

Abstract

Since thoroughly validated naturalistic affective German speech stimulus databases are rare, we present here a novel validated database of speech sequences assembled with the purpose of emotion induction. The database comprises 37 audio speech sequences with a total duration of 92 minutes for the induction of positive, neutral, and negative emotion: comedian shows intending to elicit humorous and amusing feelings, weather forecasts, and arguments between couples and relatives from movies or television series. Multiple continuous and discrete ratings are used to validate the database to capture the time course and variabilities of valence and arousal. We analyse and quantify how well the audio sequences fulfil quality criteria of differentiation, salience/strength, and generalisability across participants. Hence, we provide a validated speech database of naturalistic scenarios suitable to investigate emotion processing and its time course with German-speaking participants. Information on using the stimulus database for research purposes can be found at the Open Science Framework project repository GAUDIE: https://osf.io/xyr6j/.

Keywords: affective reactions, emotion, arousal, valence, German auditory stimulus database, naturalistic stimuli, emotional speech, clustering, validation methods

3.2 Introduction

Emotional processing, identification, and expression all serve to facilitate interaction by understanding others' current feelings and intentions. We experience strong emotions, particularly in social scenarios, and attach high relevance and importance to social cues (e.g., Bradley, 2009; Vuilleumier, 2005). Human emotions can be communicated auditorily via sounds like speech. Research pursuing ecological validity has to employ complex, naturalistic stimuli rather than strictly controlled, fragmented, and contextually detached emotional events to investigate how the brain processes stimuli close to everyday emotional situations.

The emotional expression within speech is modulated mainly by two components. One is the prosody, which is formed by the slow modulation of acoustic parameters such as intensity, rate, and pitch (Banse & Scherer, 1996; Wilson & Wharton, 2006). The other is the lexical content of the utterance. To validly extract the communicated emotion from speech, both components need to be considered. Contradicting information from semantic content and prosody can fundamentally alter the emotion – for instance, in the case of sarcasm or irony (Pell et al., 2011). Ben-David et al. (2016) identified three phenomena when investigating the interaction between prosody and semantic content in emotional speech. First, emotional strength is reinforced for emotionally congruent stimulus presentation (supremacy of congruency). Second, one component cannot be selectively attended and separated from the other (failure of selective attention). Third, prosodic information is weighted as more important than semantic content in emotional speech processing (prosodic dominance).

While non-speech auditory stimuli can be used almost independently of the participants' native language (Belin et al., 2008; Bradley & Lang, 2007; W. Yang et al., 2018), emotional speech stimuli, utilising both prosodic and semantic components, must be language-specific. According to statistics from 2022, approximately 130 million people worldwide have German as their mother tongue or second language. Further, it is the most widely spoken mother tongue in the European Union and even one of the most widely spoken languages worldwide (Statista Research Department, 2022). Although the distribution of German speakers and the active research community in this linguistic area are substantial, there is currently a lack of carefully developed and validated naturalistic stimulus databases for emotion induction. Stimulus material of a naturalistic database should be validated regarding criteria of (i) differentiation of conditions (specificity of the induced emotion), (ii) sufficient salience, intensity, and intrusiveness, and (iii) generalisability across individual differences in order to ensure successful emotional induction (Westermann et al., 1996).

There are several approaches to develop an emotional speech database. Kanske and Kotz (2011) addressed the lexical component by choosing 120 words from the Leipzig Affective Norms for German (LANG), which were validated using the dimensions valence and arousal in written and auditory form. Another common approach is to record semantically neutral sentences or non-comprehensible vocal utterances spoken by (professional) speakers with different emotional expressions, thereby modulating only the component of prosody (Banse & Scherer, 1996; Belin et al., 2008; Burkhardt et al., 2005; Lassalle et al., 2019; Schröder, 2003). Lassalle et al. (2019), who present a comprehensive validated emotional voice database with three subsets of uttering sentences (in British English, Swedish, and Hebrew), also provide an overview of existing databases for simulated/acted emotional vocal stimuli. Although such recorded stimuli allow a satisfactory recognition of the intended emotion due to distinct acoustic features, they disregard the lexical component.

Thus, single words and even contextually detached sentences do not properly approximate the interactive and context-sensitive characteristics of social communication. Consequently, they can be rated as less ecologically valid. In naturalistic speech, the communication of emotion can be more subtle and ambiguous (e.g., by creating contradicting information when voluntarily trying to suppress the emotion in one component due to personal or cultural principles).

Stimuli reflecting naturalistic emotional speech may be created by the following two means. The first is by developing interrogative techniques aiming to engage a person in an emotional conversation, which is a self-related emotional induction. Such self-related stimuli can be considered sufficiently strong since self-attribution might increase the relevance and strength of the emotional cue. However, comparability and generalisability among individuals are then scarcely possible. The second is by presenting already existing engaging speeches addressing the audience, or conversations of others to a person, which is experiencing emotions from the listener/observer perspective. The latter approach, using the same conversations for each individual, facilitates generalisability. However, the perceived emotion by the observer might deviate from the expressed emotion of those persons involved in the conversation. In order to warrant generalisability among individuals and identify the emotion and its strength induced in the observer, it is of utmost importance to validate the emotional speech stimuli using an independent sample. Douglas-Cowie et al. (2003) provide a recent overview of available databases and evaluate them with the following components: (1) scope (i.e., number, gender, and language of speakers, description of the emotions, contextual set-up), (2) naturalness (simulated, scripted, linguistic nature), (3) context (isolated or context-based), (4) descriptors (annotation of the material), and (5) accessibility of the database.

There are several naturalistic speech databases available, making use of either movie excerpts with expressive content and broadcasted television programs, or recordings of acted or spontaneously unfolded speeches and conversations (e.g., the HEU Emotion Database; J. Chen et al., 2021; Ryerson Audio-Visual Database of Emotional Speech and Song; RAVDESS; Livingstone & Russo, 2018; Russian Acted Multimodal Affective Set; RAMAS; Perepelkina et al., 2018; Interactive Emotional Dyadic Motion Capture Database; IEMOCAP; Busso et al., 2008; eNTERFACE'05; Martin et al., 2006; Vera am Mittag Database; VAM; Grimm et al., 2008). However, most of the databases are developed with the purpose of machine-based emotion recognition (Grimm et al., 2008; Schuller et al., 2009) rather than emotion induction. Consequently, they do not fulfil required criteria of a sufficient level of salience, intensity, and intrusiveness. Thus, despite a great need in the German-speaking area, there is currently no suitable stimulus database available which is particularly developed and validated for emotion induction using naturalistic speech.

Therefore, we aim at providing such a validated naturalistic German speech database, GAUDIE (**G**erman **AUDI**tory **E**motional database), suitable for investigating emotion processing and its time course in naturalistic scenarios. Three types of audio sequences representing speech were selected to induce positive, neutral, and negative emotions. To induce positive emotions, here in the form of humorous and amusing feelings, recordings from various comedy shows were selected. Weather forecasts were selected for the induction of neutral emotion, and arguments between couples and relatives from movies or television series were selected to evoke negative emotions.

Different emotion theories and methods can be applied to measure how well an intended emotion is induced by a stimulus. Theories assuming that emotional states are discrete suggest that distinguishable (basic) emotions, anger, fear, disgust, sadness, joy, surprise (Ekman, 1992a; Izard, 2007), as well as larger selections of emotions (e.g., Ekman, 1992a; Izard, 2007), may be evoked by stimulus material. This allows us to evaluate stimuli regarding these emotional states using accuracy of categorisation as measure. Other theories describe emotions via continuous uni- or multidimensional scales; for instance, valence (i.e., the pleasantness of an event or stimulus), arousal (i.e., the activation evoked), and intensity or dominance (Bakker et al., 2014; Bradley & Lang, 1994; Russell, 1980). There are also attempts to combine the two approaches using continuous intensity ratings per discrete emotion (Plutchik, 2013; Scherer, 2005). Here, we mainly focused on a continuous multidimensional approach using valence, arousal, and dominance to validate the emotion induction, since neuroscientific findings suggest underlying neurophysiological representations of these dimensions, particularly valence and arousal, subserving all (discrete) affective states (Gerber et al., 2008; Greco et al., 2021; Posner et al., 2005, 2009).

In naturalistic scenarios, display of emotions is highly context-dependent and reveals a high variability over the course of time (between, but also within, individual sentences). Therefore, stimuli were rated not only in a post-presentation rating but also continuously during the listening to capture the variabilities of their valence and arousal. To provide a comprehensive description for each stimulus, we also include post-presentation ratings regarding a discrete emotion classification (i.e., basic emotions as well as various discrete emotions with associated intensity evaluation) and potential moderators (e.g., familiarity). We validate the database regarding the three criteria of differentiation, sufficient salience/strength, and generalisability across participants, and propose Monte Carlo simulation (MCS)-based subtractive comparisons to verify an optimal stimulus selection.

3.3 Methods

3.3.1 Stimuli

Initial stimuli were selected with the criteria that (i) there was no music in the audio sequence, (ii) the language was free of accent, and (iii) emotionally negative and positive audio sequences had high emotional intensity but were not ethically critical in content. Humorous and amusing audio sequences were chosen to elicit positive emotions, since other positive emotions, such as love or passion, might be more tailored to personality and individual experiences and require a longer time course of exposure. A pre-selection of 60 audio sequences, fulfilling the above-stated criteria (no background music, free of accent, and high in emotional intensity), was obtained via polls published in private social media groups of the students and employees of the university department. With this approach, we intended to obtain a relatively diverse initial set. Nineteen of the audio sequences were associated with negative emotions, 25 associated with positive emotions, and 16 associated with neutral emotions. Two independent raters (27 [first author] and 24 years old, both female, and German natives) rated the pre-selection using a scale from 0 (bad) to 10 (good), and provided justification for their rating. Audio sequences with an average rating below 6 were excluded. Reasons for exclusions were, among others, too short duration of the audio sequences and insufficient audio quality or speech intelligibility.

The interclass correlation between the raters examined using the Spearman rank correlation coefficient was $r_s = .92$, with a full agreement between the two raters in $48.3\ \%$ of the cases. The final stimulus pool presented in the validation study comprised a total of 37 audio sequences, covering in total duration 92 minutes, with 11 associated with negative, 11 with positive, and 15 with neutral emotions. The average of the initial suitability rating across these audio sequences was 7.37 ($SD = 1.0$, range: 6 to 10). The audio sequences were on average 148.8 seconds long ($SD = 64.12$, range: 59 to 393 seconds), and normalised to an average amplitude of -20 dBFS (decibels relative to full scale) using the *RMS* (root mean square) measure as implemented in pydub (version 0.25.1; Robert, 2011). Detailed information about each audio sequence is provided in Supplementary Table 7.1 in the Supplementary Material. All audio sequences are made available on request for scientific use in the Open Science Framework (OSF) project repository GAUDIE https://osf.io/xyr6j/.

3.3.2 Participants

Twenty-six healthy volunteers (mean age $= 24.69$ years, $SD = 3.41$, range: 19 to 33 years, 19 females, 7 males) participated in the validation study at the University of Oldenburg, Germany. Only participants reporting German as their mother tongue, no neurological diseases or psychiatric disorders, as well as no consumption of centrally effective medication or drugs were invited. Regarding their educational level, $50\ \%$ of the participants at least graduated from high school (here the German Abitur), while $38.5\ \%$ reported a bachelor's and $11.5\ \%$ a master's degree as their highest educational level.

Previous studies reported effects of personality trait differences on the neuronal responses associated with emotion processing (e.g., emotional word recognition, Ku et al., 2020; and emotional image processing, Kehoe et al., 2012), and on the relationship between valence and arousal ratings (Kuppens et al., 2017). Kanske and Kotz (2012) described the influence of depression and anxiety on valence and arousal ratings of auditory affective words in a non-clinical sample. The authors emphasised the importance of including such psychological characteristics of the rating samples when validating databases, to enable subsequent studies to evaluate the match of their sample to the original rating sample (Kanske & Kotz, 2012). Therefore, we recorded and report personality traits, subjective well-being, and anxiety to provide a comprehensive description of the sample characteristics. Participants completed questionnaires to assess their personality via the short version of the German Big Five Inventory (BFI-K; Rammstedt & John, 2005), subjective well-being via the WHO-5 (World Health Organization, 1998), and Satisfaction With Life Scale (SWLS; Janke & Glöckner-Rist, 2012), depression via a two-item inventory (Whooley et al., 1997), and anxiety via the German State-Trait Anxiety Inventory (STAI; Laux et al., 1981). Results of the personality scores, as well as scores regarding the average subjective well-being and self-reported anxiety, indicated a healthy sample and are reported in detail in the Supplementary Material (Supplementary Figure 7.1, Supplementary Figure 7.2, and Supplementary Table 7.2). The internal consistencies using Cronbach's α of the used scales are reported in the Supplementary Table 7.3. Participants gave their written informed consent according to the recommendations of the Declaration of Helsinki before participation and received monetary compensation. The study protocol was approved by the ethics committee of the University of Oldenburg, Germany (Drs. EK/2018/070).

3.3.3 Procedure

Participants listened to only half of the audio sequences to ensure that the session did not exceed two hours. In each session, the number of stimuli was balanced across conditions, and stimulus order was randomised. Each participant provided a continuous rating during and a post-presentation rating after the presentation of each audio sequence. We instructed participants to evaluate what listening to the audio sequences triggered in them personally, rather than to rate the displayed feelings of protagonists in the audio sequence (see Supplementary Material 7.3, for the written German instructions and English translation).

For the continuous rating – that is, when listening to the audio sequence – an adapted version of the Self-Assessment Manikin (SAM; Bradley & Lang, 1994) subscale of valence and arousal with a slider below the pictorial rating scale was used (see Supplementary Figure 7.3). Each audio sequence of the individual selection was presented twice to the participants. In the first presentation, valence was rated, and in the second, arousal. We chose this order because we assumed that repetition might affect the evaluation of valence more strongly than the evaluation of arousal. Moreover, at least one other sequence was presented between the two presentations. The slider position was sampled at 1 Hz, mapped on a scale between 0 and 100, and was placed at stimulus start in the middle of the scale.

The post-presentation ratings of the audio sequences, which were taken after the presentation, included multiple rating scales. After the first presentation of the audio sequence, participants selected the prevailing emotion they experienced during the presentation on an adapted version of the Geneva Emotion Wheel (GEW, version 3.0; Sacharin et al., 2012), with an intensity rating on a 6-point Likert scale (see Supplementary Figure 7.3). We removed the option to choose neutral because we wanted to obtain the direction of a neutral audio sequence with a forced choice. A neutral evaluation of the experienced emotion could still be expressed via the intensity of the chosen emotion. Next, after the GEW rating, they had to select the prevailing emotion out of the six basic emotions (joy, anger, disgust, fear, sadness, and surprise; Ekman, 1992a; Izard, 2007).

Many neurophysiological studies investigating emotional processing (e.g., Lithari et al., 2010; Posner et al., 2009; Rozenkrants & Polich, 2008), and database validation (Belin et al., 2008; Lima et al., 2013), focused only on two SAM dimensions – valence and arousal. Bakker et al. (2014), however, strongly advocate in their review the inclusion of the third SAM dimension, dominance, to comprehensively evaluate the emotional processing. We decided, thus, to include dominance to explore differentiation of the stimulus types, but only as a short post-presentation scale to save time. Hence, after the second presentation, participants were asked to rate the experienced dominance of the previously presented audio sequence via the SAM rating and how familiar they were with the audio sequence, using a slider ranging from 0 to 100, with the starting point at 50. A major rationale for including dominance only as post-presentation rating was that an assessment of one's perceived control and dominance triggered by the stimuli might be difficult and potentially biased since it is strongly affected by the experimental-based inescapable exposure. Previous studies using non-speech auditory stimuli and focusing less on the sensations induced in the raters reported that the dominance ratings mirrored the valence dimension (Bradley & Lang, 2007; W. Yang et al., 2018). Because of the differences in task instruction and stimulus type, we here expect a similar evaluation to the arousal dimension, with dominance representing the extent to which a stimulus engages and captures the listener's attention.

3.3.4 Statistical Analysis

Descriptive Statistical Analysis

In the first part, the stimulus database was analysed descriptively using (a) the SAM ratings with the subscales valence and arousal, as well as the post-presentation subscale dominance, and (b) the additional post-presentation ratings (familiarity, basic emotions, GEW emotions, and strength of the GEW emotions).

SAM Ratings. For the continuous valence and arousal ratings, we visualised each audio sequence as a time series by calculating the bootstrapped mean and its $95\ \%$ confidence interval (CI) per time point across participants. In the bootstrapped time series, we observe an orientation phase at the beginning of each audio sequence. In this orientation and response adaptation phase, participants first learned about the context and protagonists of the audio sequence, and moved the slider, initially positioned at the middle of the scale, little in all three conditions. After approximately 30 seconds, there is a shift from the average starting point of the slider towards lower valence ratings in the negative condition, and higher ratings in the positive condition. In further analyses, we, therefore, removed this time interval of orientation and response adaptation in all three conditions. Next, we computed the grand average and its $95\ \%$ CI of the mean and standard deviation for each audio sequence by first averaging over the time series within each participant and then bootstrapping with 5,000 iterations across participants. The latter step was also applied to the dominance ratings, allowing us to obtain the bootstrapped mean and its $95\ \%$ CI over participants. The grand average over audio sequences of the same condition was calculated as a reference for within-class similarity. In the next descriptive analysis using the SAM scales, we introduce a clustering approach to further examine separation and discrimination among audio sequences and conditions using fuzzy c-means (Bezdek et al., 1984; Nayak et al., 2015). We used a pre-defined number of three clusters, representing the three conditions, and the Python fuzzy c-means toolbox (version 1.6.4; Dias, 2019). The fuzzy c-means algorithm allows for a fuzzy partitioning into clusters. This algorithm assigns a probability of belonging to a certain cluster to each audio sequence. This probability of cluster affiliation can be used as a score to assess how representative a data point is of a cluster. To determine the cluster affiliation, those with the highest probability are chosen. The matching between cluster and condition was determined by the majority of audio sequences belonging to one experimental condition within the respective cluster. From the obtained probability scores of the ratings per audio sequence, we calculated its mean and $95\ \%$ CI for each cluster affiliation via bootstrapping with 5,000 iterations. With good separation, the lower bootstrapped CI of the probability associated with the true cluster affiliation (i.e., the condition of the respective audio sequence) should not overlap with the CI of the probabilities for the other clusters (mainly the neighbouring neutral condition).

In a final analysis of the SAM ratings, we were interested in increasing the differences between ratings of the emotional audio sequences and those of the neighbouring condition (i.e., the neutral condition), by choosing only particularly distinct audio sequences. These are audio sequences revealing a large difference from those of the neighbouring neutral condition and small difference from audio sequences of the same emotional condition. To evaluate how well each emotional audio sequence was discriminated from neutral

audio sequences in their SAM ratings, we performed repeated pairwise subtractive comparisons via MCS (e.g., Ernst, 2004) per subscale. In the first step, all possible combination pairs of emotional and neutral audio sequences were generated (total combination pairs of $N = 330$ with $n_{negative-neutral} = 165$ and $n_{positive-neutral} = 165$). For each combination pair, we created a vector of 5,000 randomly selected ratings of the emotional audio sequence and subtracted it from a vector of 5,000 randomly selected ratings of the neutral audio sequence, thereby ignoring the within-subject factor. Based on the hypothetically assumed relationship of both audio sequences, we identified combination pairs revealing less distinct separation. For this purpose, we calculate the percentage of occasions per combination pair in which the hypothetically assumed relationship of both audio sequences was violated. For the valence ratings, we assumed increased values from negative to neutral to positive. For the arousal and dominance ratings, we assumed higher values for emotional (negative and positive) compared to neutral audio sequences. We visualised those pairs with a probability of misclassification above $20\ \%$. In application cases where a particularly distinct pool of audio sequences per condition is required, this analysis allows justification to select a subset of audio sequences.

Additional Post-Presentation Ratings. In the second part of the descriptive analysis, the database was analysed regarding the additional post-presentation rating scales, comprising familiarity, basic emotions, GEW emotions, and strength of the GEW emotions.

Frequencies of the GEW and basic emotions ratings were normalised between 0 and 1, and visualised using radar charts per condition and each audio sequence. We computed four additional scores from the GEW and basic emotions ratings, describing whether the participant rated the audio sequence negatively or positively. Therefore, we first classified whether the chosen emotion was associated positively or negatively. Regarding the GEW ratings, the emotions interest, joy, pleasure, satisfaction, amusement, compassion, pride, and relief were classified as positive, whereas fear, disgust, hatred, anger, disappointment, sadness, contempt, shame, guilt, and regret were classified as negative. Regarding the basic emotions, joy and surprise were classified as positive and sadness, fear, disgust, and anger as negative. This resulted in four additional scores (i.e., negative basic emotion, positive basic emotion, negative GEW emotion, and positive GEW emotion). In the next step, only the GEW emotion scores were weighted by the strength of the perceived response using the GEW intensity rating.

Inferential Statistical Analysis

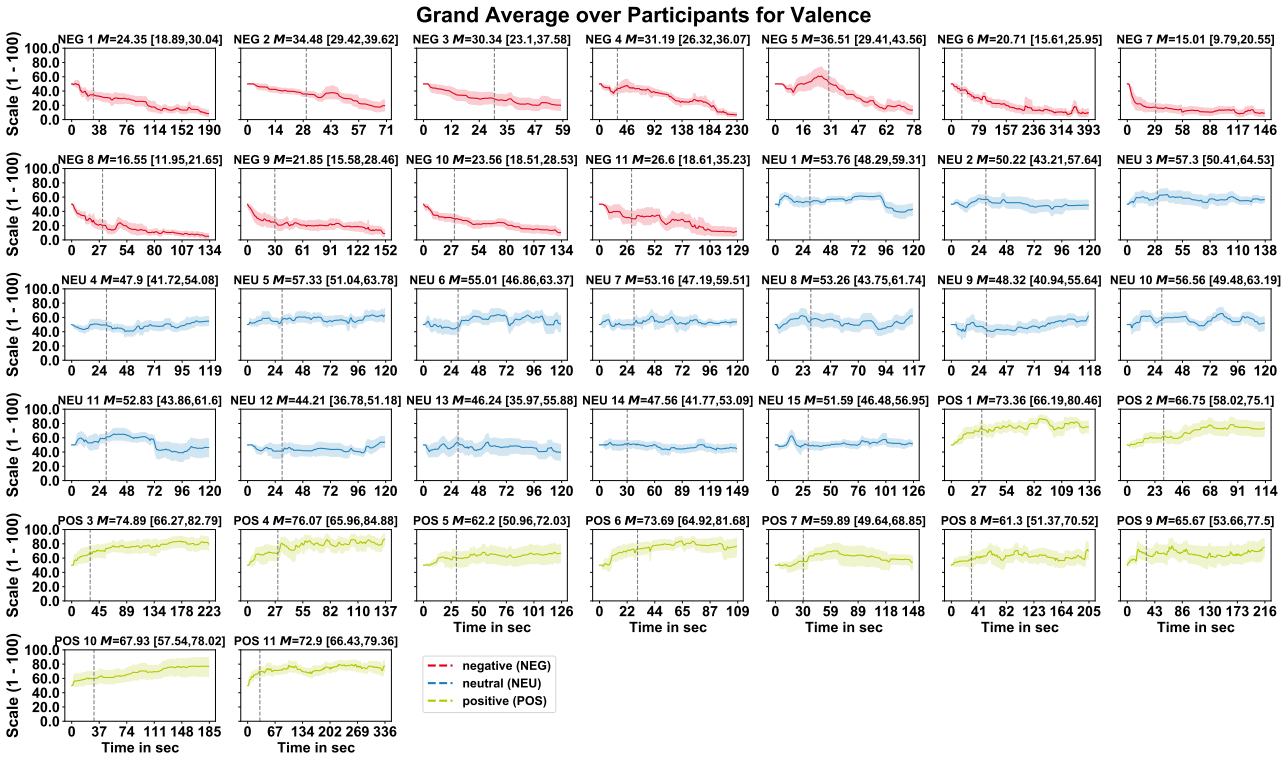
Finally, the three stimulus conditions were statistically compared regarding their SAM ratings and additional post-presentation ratings (i.e., familiarity, basic emotions, negative GEW emotion, positive GEW emotion, GEW strength/intensity) using repeated-measures analyses of variance (rmANOVAs; toolbox pingouin; version 0.5.1; Vallat, 2018) with the factor condition (negative vs. neutral vs. positive). Post-hoc comparisons between conditions were performed by computing the difference participant-wise, as well as the Bonferroni-corrected CI for each contrast via bootstrapping. Contrasts, where the CI does not include zero, can be considered as significant (Cumming & Finch, 2005).

3.4 Results

3.4.1 Descriptive and Inferential Analysis Using the SAM Ratings

Figure 3.1

Bootstrapped Grand Averages for the Continuous Valence Ratings Among Participants per Audio Sequence Over the Time Course of Listening



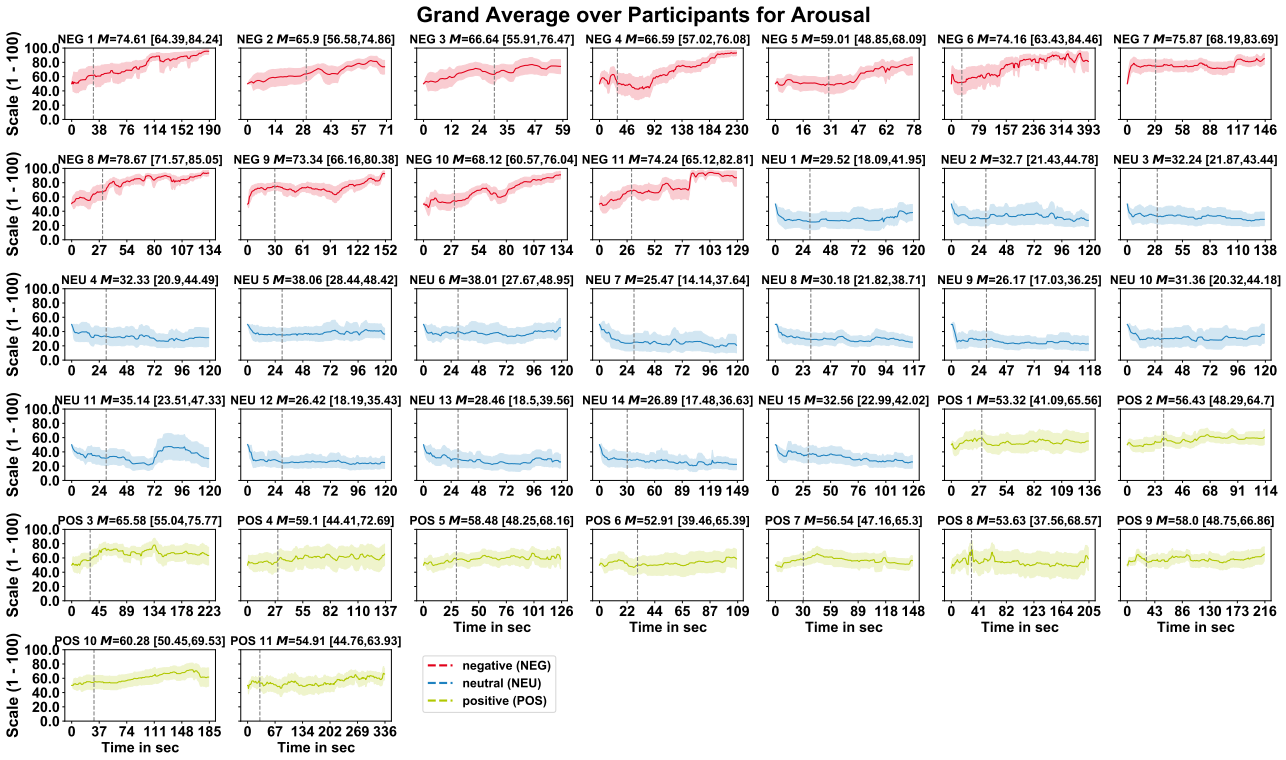
Note. Shaded areas represent the bootstrapped $95\ \%$ confidence interval (CI) of the mean deviation over participants. The mean (M) over the time course with its $95\ \%$ CI are provided in brackets in the subtitles. The dashed vertical line indicates the initial 30-second adaptation phase.

To investigate the variability of the ratings over the time course of listening, Figures 3.1 and 3.2 illustrate the bootstrapped mean and its $95\ \%$ CI per time point across participants for each audio sequence for the continuous valence and arousal ratings. After the orientation and response adaptation phase, lower valence ratings in the negative condition and higher ratings in the positive condition can be observed. For the continuous arousal ratings, we observed lower values for the neutral audio sequences compared to the negative and positive conditions, as well as lower arousal values for the positive when compared to the negative condition. Nevertheless, we observe dynamic variation in the temporal course of listening for both continuous ratings, which illustrates the importance to keep the evaluation for dynamic, complex stimuli at the level of individual time

points. We also examined whether the position within a session affected the continuous rating. No correlation between ratings and their position was observed for valence ($r_p = -.03; p = .471$) and arousal ($r_p = .002; p = .972$).

Figure 3.2

Bootstrapped Grand Averages for the Continuous Arousal Ratings Among Participants per Audio Sequence Over the Time Course of Listening



Note. Shaded areas represent the bootstrapped $95\ \%$ confidence interval (CI) of the mean deviation over participants. The mean (M) over the time course with its $95\ \%$ CI are provided in brackets in the subtitles. The dashed vertical line indicates the initial 30-second adaptation phase.

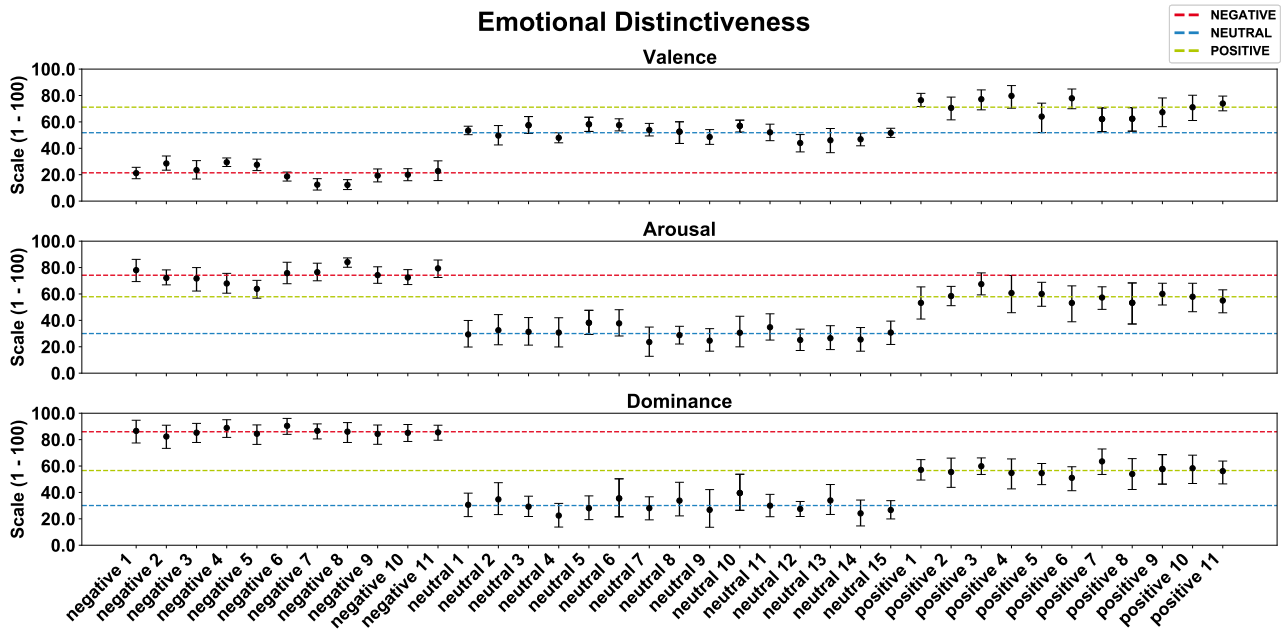
To examine the strength of the induced emotional evaluation, generalisability across participants, and differentiation to audio sequences of the neighbouring conditions, Figure 3.3 shows for each stimulus the bootstrapped grand average over time and participants of the continuous valence and arousal ratings, as well as the grand average over participants of the post-presentation dominance ratings, excluding the initial 30-second adaptation phase. Error bars indicate the $95\ \%$ CI of the grand average. The horizontal dashed lines represent the grand average over all stimuli of each emotion condition.

Generally, the SAM ratings are similar within conditions and exhibit good separation between conditions. However, the CIs of the valence ratings for the positive audio sequences 5, 7, and 8 slightly touch the grand average of the neutral emotion condition, suggesting that these stimuli may be rated neutral by some participants.

Similarly, we observed for arousal ratings slight overlaps of the CIs of the audio sequences negative 5, positive 3, and positive 4 with the grand average of the opposite emotional condition. Regarding the post-presentation dominance ratings, we observed good separation between all three conditions, with the highest ratings for the negative audio sequences, followed by positive sequences, and the lowest ratings for neutral audio sequences. Detailed information on the mean and CI per audio sequence and rating is provided in Supplementary Table 7.4.

Figure 3.3

Bootstrapped Average Valence, Arousal, and Dominance Ratings, as Well as $95\ \%$ *Confidence Interval (CI) per Audio Sequence*



Note. CI: confidence interval. Dashed lines represent the mean overall stimuli in each emotion condition.

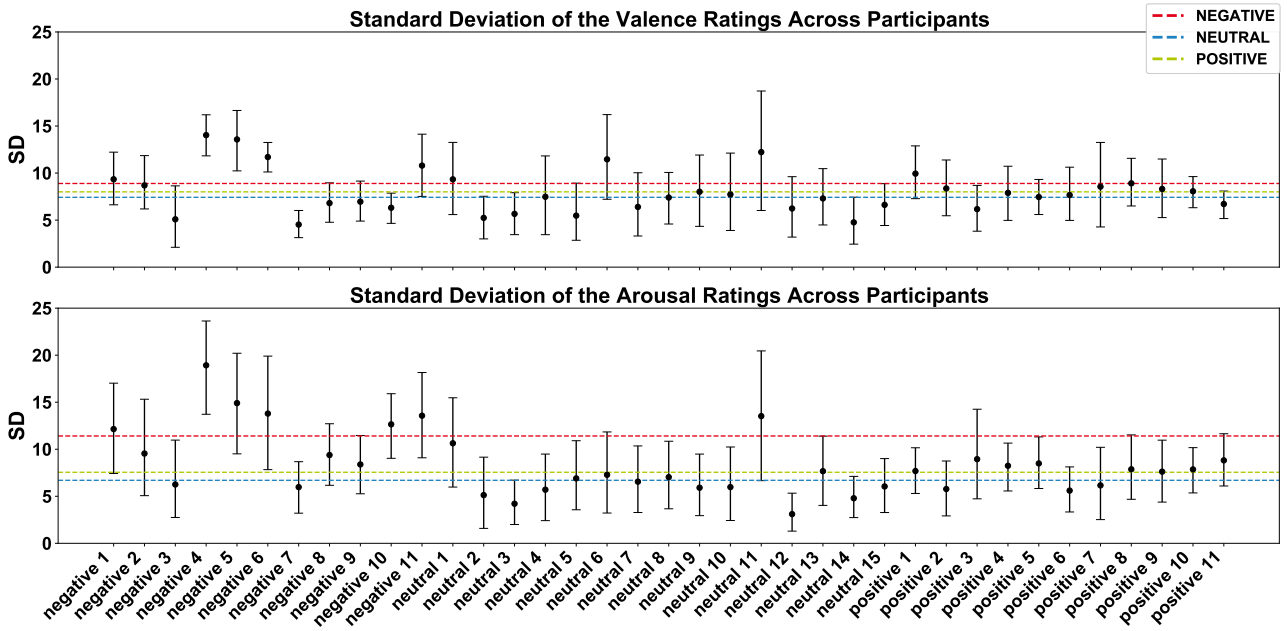
Results of the rmANOVA revealed strong significant differences of the average ratings per condition for all three SAM scales, with $F_{(2,\ 50)} = 160.22$, $p < .001$, $\eta_p^2 = 0.87$ for the valence, $F_{(2,\ 50)} = 71.73$, $p < .001$, $\eta_p^2 = 0.74$ for the arousal, and $F_{(2,\ 50)} = 134.41$, $p < .001$, $\eta_p^2 = 0.84$ for the dominance ratings. Bootstrapped mean differences and Bonferroni-corrected 2.5^{th} and 97.5^{th} CIs of the mean per contrast are provided in Supplementary Table 7.12.

The consistency of ratings within each audio sequence is indicated by the standard deviation of the continuous ratings over time. Figure 3.4 provides the bootstrapped average standard deviation and respective $95\ \%$ CI of the valence and arousal ratings for each stimulus across participants (see also Supplementary Table 7.5). For the valence ratings, we observed comparable variations within the ratings over time across conditions, with some outliers revealing larger variability (negative 4 and 5, as well as neutral 6 and 11), but no statistically significant difference between the conditions, $F_{(2,\ 50)} = 1.88$, $p = .163$, $\eta_p^2 = .07$ (see

Supplementary Table 7.12). The standard deviation of the arousal ratings is significantly higher for the negative condition than for the neutral and positive condition, indicating that the arousal ratings of the negative stimuli are less consistent over the time course of the audio sequence, with $F_{(2,\ 50)} = 15.20, p < .001, \eta_p^2 = .38$ (see Supplementary Table 7.12).

Figure 3.4

Bootstrapped Average Standard Deviation (SD) of the Valence and Arousal Ratings as Well as Its $95\ \%$ Confidence Interval (CI) per Audio Sequence

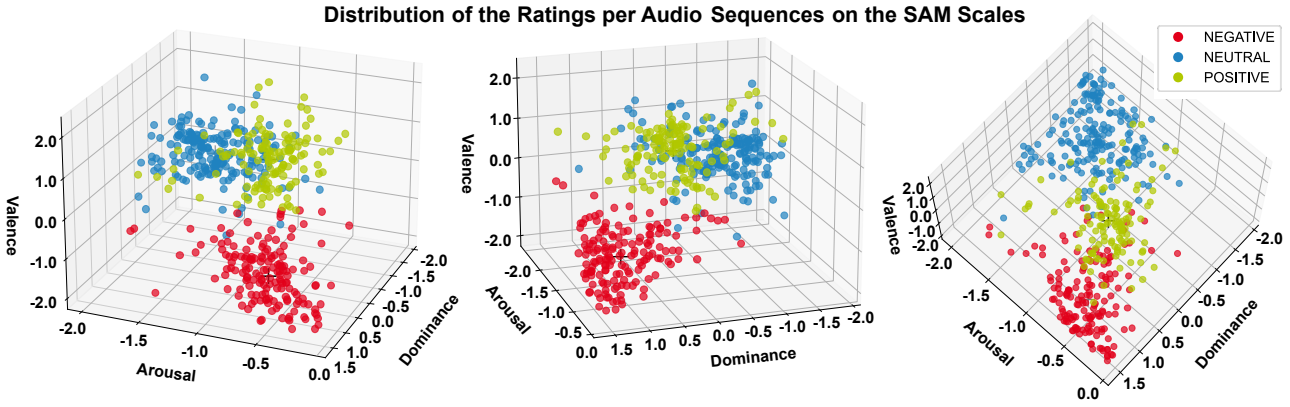


Note. CI: confidence interval. Dashed lines represent the mean overall stimuli in each emotion condition.

Next, we investigated the separability and clustering of the stimuli in the full SAM space with its three dimensions instead of analysing each dimension separately. To quantify the separation of the audio sequences of the three conditions, we performed an unsupervised fuzzy c-means clustering with three clusters and the SAM ratings as feature set. The cluster assignment had a $93.02\ \%$ correspondence with the a-priori defined conditions, indicating that the different stimuli form the expected clusters on the SAM ratings, and are well separated between clusters. The matching between cluster and condition was determined by the majority of audio sequences belonging to one experimental condition within the respective cluster. Figure 3.5 illustrates the distribution of the clustered data in a three-dimensional space where each axis represents a SAM scale, and each point a stimulus rating from one participant. The data clearly fall into three clusters representing the three conditions. The red point cloud, representing the negative condition, appears to be separated by the largest margin from the other clusters. The positive (green) and neutral (blue) clusters are less separated.

Figure 3.5

Distribution of the Clustered Z-Score Normalised Data on the Three SAM Scales of Valence, Arousal, and Dominance



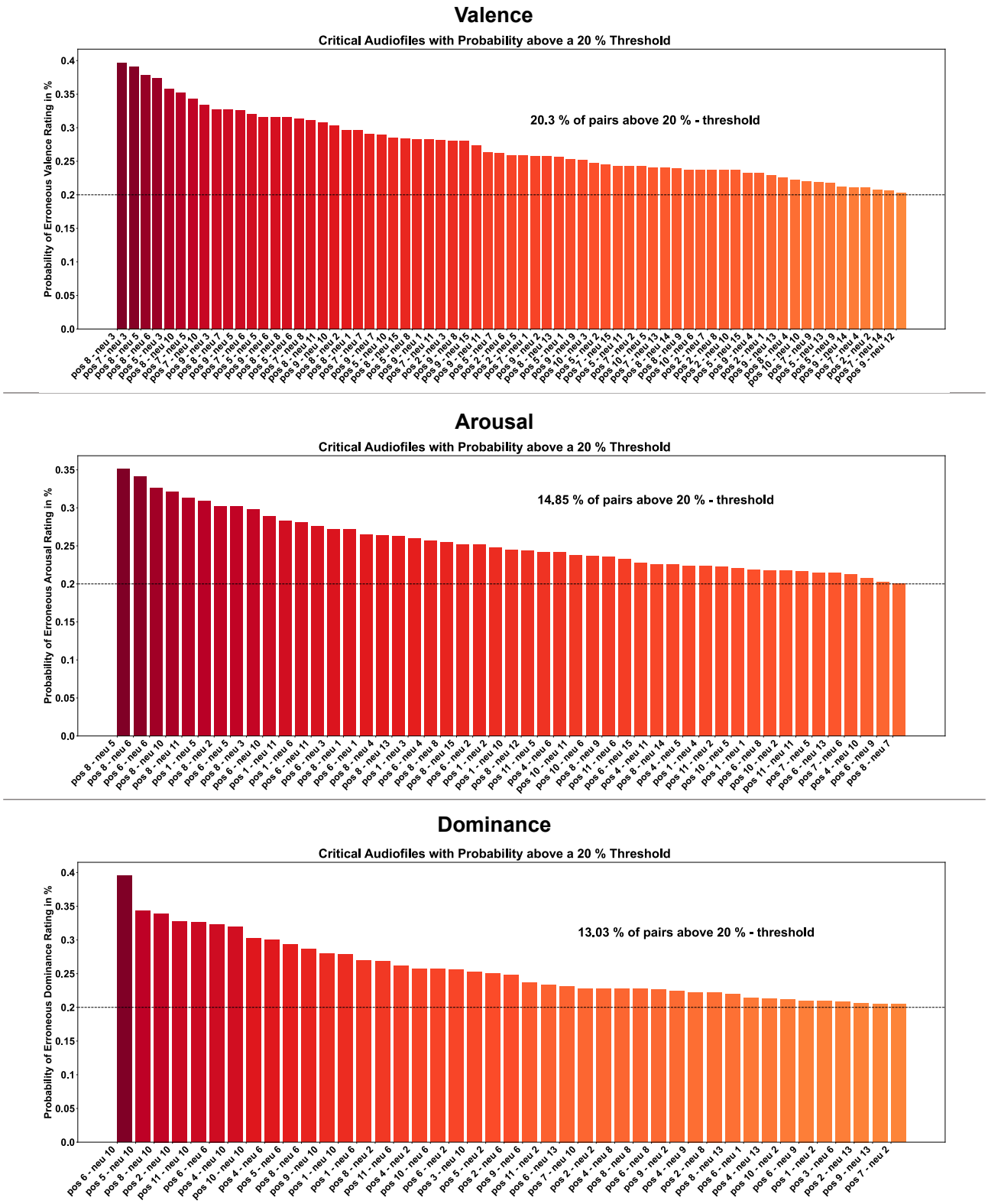
Note. Each data point represents a clustered audio sequence rating of one participant. Colours represent the cluster label.

We further analysed the separability of stimuli that were misclassified using the probability of cluster affiliation per cluster. Fuzzy c-means allows the computation of the probability of the assignment to each of the three clusters. We calculated the bootstrapped average probabilities and CIs across participants per cluster affiliation of audio sequences, which were at least once (i.e., for at least one participant) wrongly assigned based on the fuzzy c-means clustering. For each audio sequence, we marked the lower CI of the average probability of the true cluster affiliation across participants as a threshold (see Supplementary Figure 7.4). In case the CI of the average probability for belonging to one of the other two clusters overlaps with or touches this threshold, the audio sequence is less clearly differentiated from neighbouring clusters. Using this logic, only the positive 8 audio sequence is identified as less separable from one other cluster (i.e., the neutral condition) in the SAM rating space (probability for the neutral cluster: $M = 0.29\ [0.17,\ 0.42]$; probability for the positive cluster: $M = 0.56\ [0.40,\ 0.72]$; see Supplementary Figure 7.4 and Supplementary Table 7.6, for further values of the bootstrapped means and CIs). This indicates that even when the stimulus ratings of single participants lead to wrong emotion classifications, the stimuli are well separated at the level of the average SAM ratings in the participant population.

Finally, we investigated which audio sequences could be removed from the stimulus set if only a subset of stimuli is required in an experiment. In our approach, we simulate the discriminability of stimulus pairs on the SAM ratings using a Monte Carlo Simulation (see Methods) to identify audio sequence pairs particularly suitable for removal due to lower discriminability.

Figure 3.6

Repeated Pairwise Subtractive Comparisons via a Monte Carlo Simulation (MCS) With 5,000 Iterations



Note. The depicted stimulus pairs have a probability of erroneous evaluation above $20\ \%$. Dashed lines represent the pre-defined threshold of 0.2, indicating a $20\ \%$ probability of an erroneous rating relationship between the combination pair. Pos: positive, neu: neutral, neg: negative.

In Figure 3.6, we visualise those stimulus pairs which revealed erroneous rating relationships with a probability of at least $20\ \%$ in the simulation, separately for each SAM scale. This rather complex listing of misclassification rates among stimulus pairs can be broken down by analysing which stimulus pairs were misclassified with a probability above $20\ \%$.

For the valence ratings (Figure 3.6, upper row), $20.3\ \%$ of all combination pairs revealed a probability of misclassification above the threshold of $20\ \%$. The top four audio sequences with the highest probabilities of misclassification were positive 7, positive 8, neutral 3, and neutral 5, and were, therefore, identified as suitable candidates for removal if the valence scale is of importance. For the arousal ratings, we observed above-threshold misclassification in $14.85\ \%$ of the combination pairs, with positive 8, neutral 6, neutral 5, and positive 6 being the top four with the most repetitions and highest probabilities (Figure 3.6, middle row). For the dominance ratings, misclassification above the $20\ \%$ threshold occurred in $13.03\ \%$ of the combination pairs. The top four audio sequences with the highest probabilities of misclassification were neutral 10, neutral 6, neutral 2, and positive 6, indicating them as suitable candidates for removal (Figure 3.6, lower row).

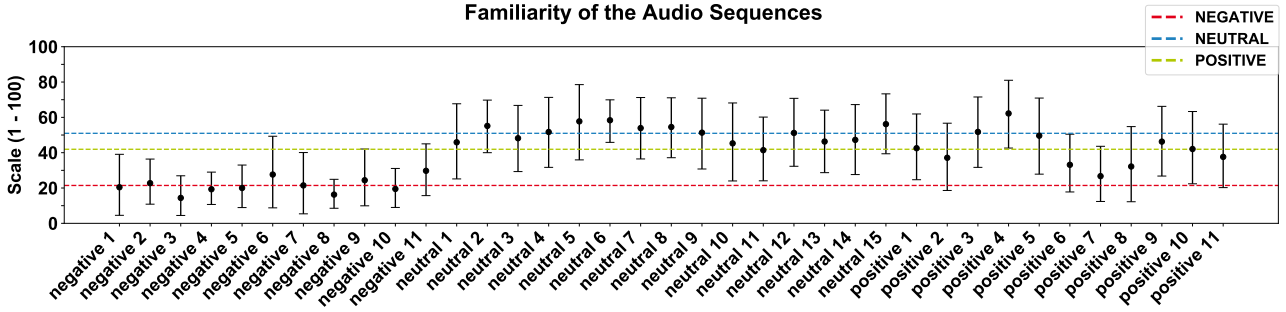
3.4.2 Descriptive and Inferential Analysis Using Additional Post-Presentation Ratings

Since familiarity of a stimulus might affect other ratings, we analysed how familiar the participants were with the audio sequences. On average, the highest ratings were assigned to stimuli of the neutral condition, followed by stimuli of the positive emotion condition. The negative emotion condition received on average the lowest familiarity ratings (see Figure 3.7 and Supplementary Table 7.7), and differed significantly from average ratings of neutral and positive stimuli, $F_{(2,\ 50)} = 13.89$, $p < .001$, $\eta_p^2 = .36$ (bootstrapped contrasts with negative – neutral: $M = -28.87\ [-39.62,\ -18.20]$; negative – positive $M = -21.01\ [-33.94,\ -8.38]$; but positive – neutral (n.s.): $M = -7.85\ [-23.56,\ 8.09]$); see Supplementary Table 7.12). To examine whether familiarity of the audio sequence influences its SAM ratings, we correlated the SAM ratings with familiarity per condition (see Supplementary Table 7.8). None of the correlations was significant. We observed a non-significant trend towards a positive correlation between familiarity and valence ratings only for the positive stimuli.

We now turn to the analysis of the discrete emotion ratings, the GEW and the basic emotions classes. Figure 3.8 shows the distribution and frequencies of chosen emotions for the basic emotions (upper) and GEW ratings (lower), separately for the three stimulus types. We observe a wider distribution of basic emotions chosen for the negative stimuli, which distributes over fear, anger, sadness, and disgust. Similarly, the GEW ratings scatter over a wide range of predominantly negative emotions, including anger, contempt, disgust, fear, disappointment, shame, and sadness, but also compassion. For the positive and neutral stimuli, we observe less variations in the chosen emotions. Among the basic emotions, mainly joy and surprise were selected, with little differentiation between stimulus types. However, positive and neutral stimuli separated on the GEW scale. There, amusement was most often chosen for the positive, and interest for the neutral condition. The distributions and frequencies per stimulus are provided in the Supplementary Figure 7.5 to Supplementary Figure 7.10.

Figure 3.7

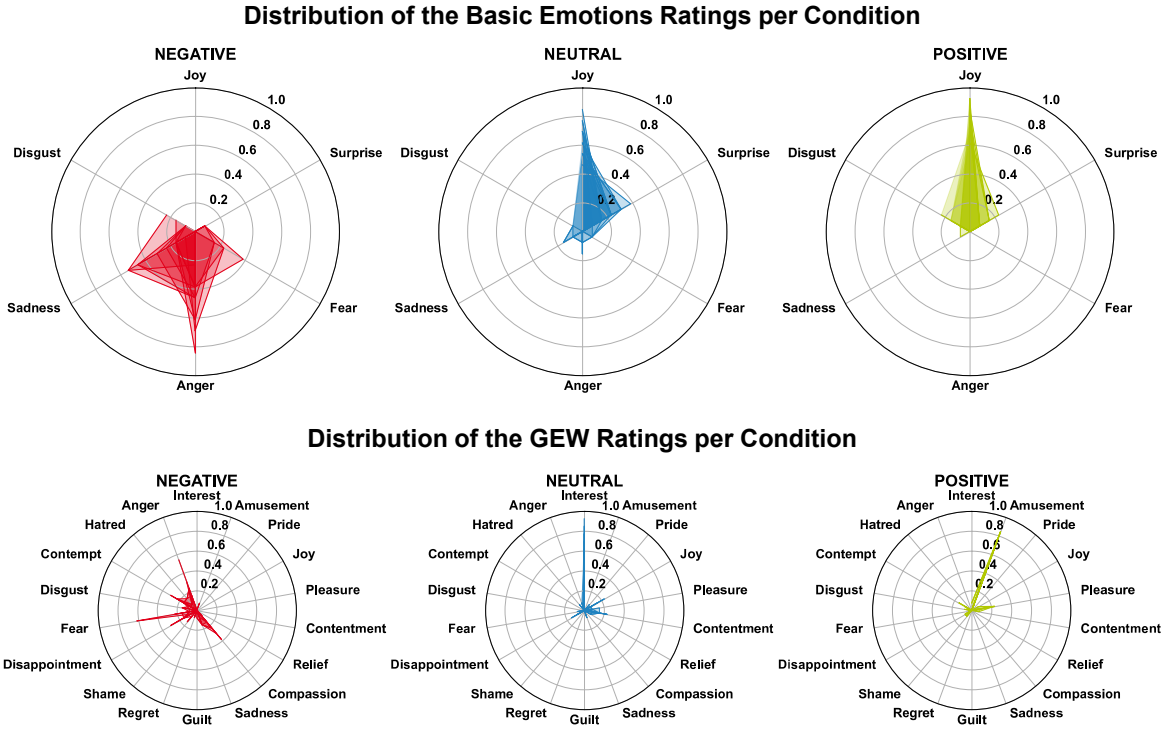
Bootstrapped Average Familiarity and Its $95\ \%$ *Confidence Interval (CI) per Audio Sequence*



Note. Dashed lines represent the grand average within the respective condition.

Figure 3.8

Frequencies of the Normalised Basic and GEW Emotions Ratings per Condition



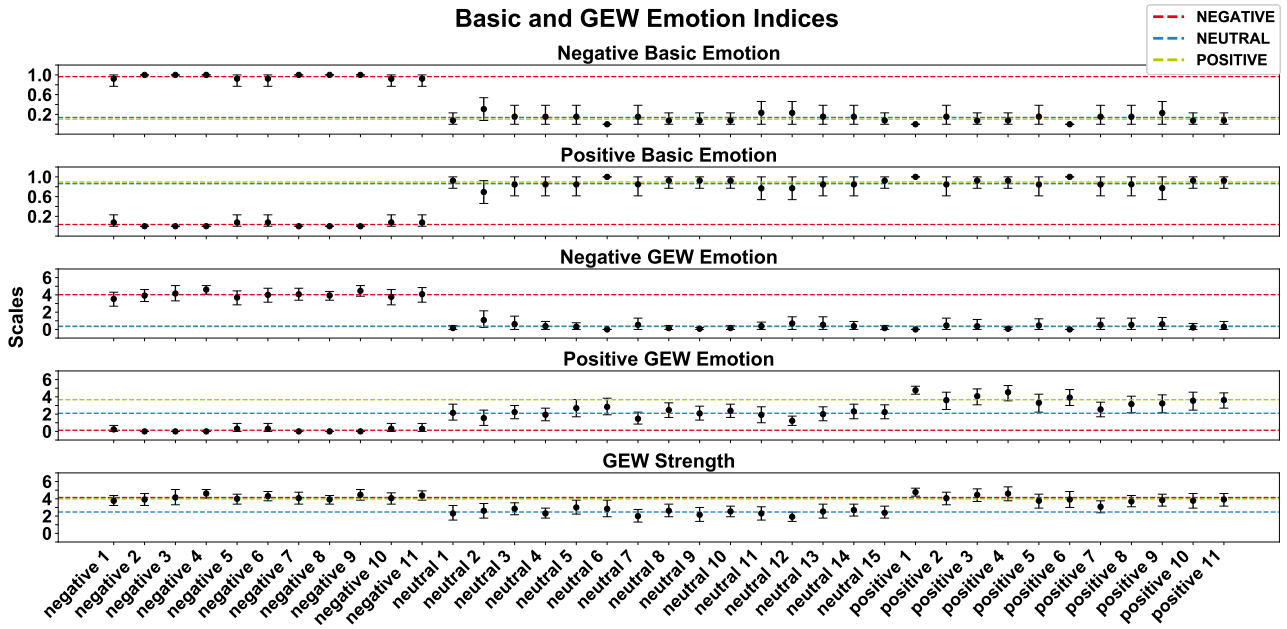
Note. Scale ranging from 0 to 1. Each layer represents the ratings of one stimulus.

Figure 3.9 shows the basic emotions and the GEW ratings averaged across participants, separately for each stimulus. In order to reduce the complexity, we combined the ratings, separately for the negative and

positive emotions, in each discrete emotion system. The negative and positive basic emotions ratings clearly differentiated the negative from the positive and neutral stimuli, with $F_{(2,\ 50)} = 170.50, p < .001, \eta_p^2 = .87$. However, there was no significant difference between the average ratings of the positive and neutral stimuli, with a bootstrapped mean of $M = 0.03$ and Bonferroni-corrected CIs including the zero; $95\ \%\ CI\ [-0.12,\ 0.16]$. We observe a similar strong differentiation of negative from positive and neutral stimuli in the GEW ratings, with $F_{(2,\ 50)} = 231.70, p < .001, \eta_p^2 = .90$, for the negative, and $F_{(2,\ 50)} = 70.21, p < .001, \eta_p^2 = .74$, for the positive GEW ratings. We observed a significant stepwise difference only for the positive GEW score, with the lowest values in the negative condition, followed by the neutral condition, and the highest values in the positive condition (bootstrapped contrasts with negative – neutral: $M = -1.95\ [-2.52,\ -1.39]$; positive – neutral: $M = 1.6\ [0.66,\ 2.38]$; negative – positive: $M = -3.55\ [-4.11,\ -2.9]$; see Supplementary Table 7.12).

Figure 3.9

Bootstrapped Average of the Basic and GEW Emotions Ratings With Its $95\ \%$ *Confidence Interval (CI) per Audio Sequence*



Note. Dashed lines represent the grand average within the stimulus classes.

The results suggest that the GEW ratings allowed for somewhat better differentiation between the neutral and positive conditions compared with the basic emotions ratings. The GEW strength rating clearly differs between neutral and emotional stimuli, $F_{(2,\ 50)} = 38.60, p < .001, \eta_p^2 = .61$, with lower ratings for the neutral than the negative and positive stimuli. However, no difference was found between positive and negative stimuli (bootstrapped difference negative – positive: $M = 0.13\ [-0.26,\ 0.54]$; see Supplementary

Table 7.12). Numerical values of the bootstrapped means and CIs per audio sequence for the basic emotions and GEW ratings are provided in Supplementary Table 7.9, 7.10, and 7.11.

3.5 Discussion

We here developed, validated, and characterised a novel German database of naturalistic and continuously rated auditory stimuli comprising emotionally neutral, positive, and negative speech sequences. The database aims to fill a gap in the existing databases of emotion-inducing stimuli. We demonstrate that the stimuli fulfil the criteria of a sufficient level of salience, intensity, and intrusiveness, and are well separated on the valence-arousal-dominance system of emotions.

The analyses chosen in this work aim to provide (a) multiple annotations per audio sequence, which can also be used for further validations, or for choosing a specific selection of audio sequences, and (b) an estimation of robustness across evaluation methods and participants. The rich annotations per audio sequence are a particular advantage of the database, with both post-presentation ratings and continuous annotations for valence and arousal. These annotations allow, among others, further correlation analyses with peripheral and neurophysiological reactions evoked by the processing of the audio sequences.

The participants' post-presentation basic emotions classifications did not differentiate between neutral and positive audio sequences; only negative stimuli were categorised with different emotions. Consequently, a categorical assessment based on the basic emotions may not be sufficient to distinguish between emotionally positive and neutral audio sequences. However, the GEW ratings, featuring a more fine-grained differentiation of emotion categories and a continuous intensity scale, distinguished between positive and neutral audio sequences, especially through the intensity ratings. Interestingly, we observed a difference in the familiarity post-presentation ratings for the negative compared to the positive and neutral conditions. The higher familiarity ratings for the neutral audio sequences potentially indicate that participants are more familiar with the general type of audio sequence – that is, weather forecasts – than with the specific sequence. Importantly, familiarity did not affect the SAM ratings significantly.

One limitation of the neutral audio sequence is the higher information density conveyed within the sequence. Although the conveyed information is potentially irrelevant and therefore less salient, the processing complexity increases.

A further limitation is the young participant sample with a high proportion of female university students recruited for validation. Since the validation using multiple scales and continuous ratings was quite time-intensive, only half of the database was rated by a participant. Although our results revealed rather good consistency across participants in the differentiation between stimulus types, ratings might still have been influenced by age or potential gender differences. However, the participant sample recruited for validation fulfilled specific inclusion criteria (e.g., young age below 35 years and no diagnosed psychological disorders) to approximate samples often used in neurophysiological studies. It is therefore important to note that the results of the validation may not be representative of the whole population but rather limited to a subsample of the population matching the characteristics of the validation sample. The characterisation of the stimuli can be

extended in the future. For example, the relationship between the ratings and auditory features such as pitch, timbre, and loudness, as well as speech-related features (e.g., spectral and temporal modulation features; Boos et al., 2021; Santoro et al., 2014) of the sequence could be a valuable extension. Moreover, information about the frequencies of the words used in the audio sequences, as well as speed of speech, could further widen the field of research in which the GAUDIE database is applicable.

3.6 Conclusion

We hereby provide a validated and richly annotated naturalistic affective German speech stimulus database named GAUDIE, which can be used to induce positive, neutral, and negative emotion. The multiple continuous and discrete ratings can be further used to investigate influencing auditory features, as well as the relationship between behavioural emotional evaluations and evoked (neuro-)physiological reaction, when listening to the audio sequence. The suggested audio sequences are accessible on request for further research (https://osf.io/xyr6j/), and fulfil criteria of differentiation, sufficient salience/strength, and generalisability. We hope that the stimulus set is useful for the research community and want to encourage researchers to add further annotations, which can be linked with the audio sequences.

Funding

Open Access funding enabled and organised by Projekt DEAL. The research was supported by the Fraunhofer Gesellschaft with the scholarship “Fraunhofer TALENTA”, Ministry of Economic Affairs, Labour, and Tourism Baden-Wuerttemberg in the project “KI-Fortschrittszentrum Lernende Systeme und Kognitive Robotik”, and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC 2177/1—Project ID 390895286.

Declaration of Competing Interests

The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this article. The study protocol was approved by the ethics committee of the Carl von Ossietzky University, Oldenburg, Germany (Ref. EK/2018/070).

Data and Code Availability

The Supplementary Material is accessible online in the OSF project repository GAUDIE https://osf.io/xyr6j/. The datasets generated and analysed during the current study are available upon request via the OSF project repository GAUDIE https://osf.io/xyr6j/. The GitHub repository with the code used to analyse the current study is also available in the OSF project repository GAUDIE https://osf.io/xyr6j/.

Acknowledgements

We would like to thank Jana Willemsen for her support in the data collection.

References Study II

Bakker, I., van der Voordt, T., Vink, P., & de Boon, J. (2014). Pleasure, arousal, dominance: Mehrabian and Russell revisited. *Current Psychology*, *33*(3), 405–421. https://doi.org/10.1007/s12144-014-9219-4

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. https://doi.org/10.1037/0022-3514.70.3.614

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal affective voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531–539. https://doi.org/10.3758/brm.40.2.531

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., & Lieshout, P. H. H. M. (2016). Prosody and semantics are separate but not separable channels in the perception of emotional speech: Test for rating of emotions in speech. *Journal of Speech, Language, and Hearing Research*, *59*(6), 1441–1454. https://doi.org/10.1044/2015_JSLHR-H-14-0323

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, *10*, 191–203. https://doi.org/10.1016/0098-3004(84)90020-7

Boos, M., Lücke, J., & Rieger, J. W. (2021). Generalizable dimensions of human cortical auditory processing of speech in natural soundscapes: A data-driven ultra high field fMRI approach. *NeuroImage*, *237*, 118106. https://doi.org/10.1016/j.neuroimage.2021.118106

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Bradley, M. M., & Lang, P. J. (2007). *The international affective digitized sounds (2nd edition; IADS-2): Affective ratings of sounds and instruction manual* (Technical report B-3). University of Florida.

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., Weiss, B., et al. (2005). A database of German emotional speech. *Interspeech*, *5*, 1517–1520. https://doi.org/10.21437/Interspeech.2005-446

Busso, C., Bulut, M., Lee, C.-C., Kazemzadeh, A., Mower, E., Kim, S., Chang, J. N., Lee, S., & Narayanan, S. S. (2008). IEMOCAP: Interactive emotional dyadic motion capture database. *Language Resources and Evaluation*, *42*(4), 335–359. https://doi.org/10.1007/s10579-008-9076-6

Chen, J., Wang, C., Wang, K., Yin, C., Zhao, C., Xu, T., Zhang, X., Huang, Z., Liu, M., & Yang, T. (2021). HEU emotion: A large-scale database for multimodal emotion recognition in the wild. *Neural Computing and Applications*, *33*(14), 8669–8685. https://doi.org/10.1007/s00521-020-05616-w

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *The American Psychologist*, *60*(2), 170–180. https://doi.org/10.1037/0003-066X.60.2.170

Dias, M. L. D. (2019). Fuzzy-c-means: An implementation of fuzzy c-means clustering algorithm [version 1.6.4; retrieved April 10, 2023]. https://git.io/fuzzy-c-means

Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech: Towards a new generation of databases. *Speech Communication*, *40*(1-2), 33–60. https://doi.org/10.1016/s0167-6393(02)00070-5

Ekman, P. (1992a). Are there basic emotions? *Psychological Review*, *99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Ernst, M. D. (2004). Permutation methods: A basis for exact inference. *Statistical Science*, *19*(4), 676–685. https://doi.org/10.1214/088342304000000396

Gerber, A. J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J., & Peterson, B. S. (2008). An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia*, *46*(8), 2129–2139. https://doi.org/10.1016/j.neuropsychologia.2008.02.032

Greco, A., Valenza, G., & Scilingo, E. P. (2021). Brain dynamics during arousal-dependent pleasant/unpleasant visual elicitation: An electroencephalographic study on the circumplex model of affect. *IEEE Transactions on Affective Computing*, *12*(2), 417–428. https://doi.org/10.1109/TAFFC.2018.2879343

Grimm, M., Kroschel, K., & Narayanan, S. (2008). The Vera am Mittag German audio-visual emotional speech database. *IEEE International Conference on Multimedia and Expo (ICME)*, 865–868. https://doi.org/10.1109/icme.2008.4607572

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *2*(3), 260–280. https://doi.org/10.1111/j.1745-6916.2007.00044.x

Janke, S., & Glöckner-Rist, A. (2012). *Deutsche Version der Satisfaction with Life Scale* (SWLS; Manual). ZIS – GESIS Leibniz Institute for the Social Sciences. https://doi.org/10.6102/zis147

Kanske, P., & Kotz, S. A. (2011). Cross-modal validation of the Leipzig affective norms for German (LANG). *Behavior Research Methods*, *43*(2), 409–413. https://doi.org/10.3758/s13428-010-0048-6

Kanske, P., & Kotz, S. A. (2012). Auditory affective norms for German: Testing the influence of depression and anxiety on valence and arousal ratings. *PLOS ONE*, *7*(1), e30086. https://doi.org/10.1371/journal.pone.0030086

Kehoe, E. G., Toomey, J. M., Balsters, J. H., & Bokde, A. L. W. (2012). Personality modulates the effects of emotional arousal and valence on brain activation. *Social Cognitive and Affective Neuroscience*, *7*(7), 858–870. https://doi.org/10.1093/scan/nsr059

Ku, L.-C., Chan, S.-H., & Lai, V. T. (2020). Personality traits and emotional word recognition: An ERP study. *Cognitive, Affective and Behavioral Neuroscience*, *20*(2), 371–386. https://doi.org/10.3758/s13415-020-00774-9

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The relation between valence and arousal in subjective experience varies with personality and culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

Lassalle, A., Pigat, D., O'Reilly, H., Berggren, S., Fridenson-Hayo, S., Tal, S., Elfström, S., Råde, A., Golan, O., Bölte, S., Baron-Cohen, S., & Lundqvist, D. (2019). The EU-emotion voice database. *Behavior Research Methods*, *51*(2), 493–506. https://doi.org/10.3758/s13428-018-1048-1

Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981). *Das State–Trait–Angstinventar (STAI): Theoretische Grundlagen und Handanweisung* (Manual). Beltz Test GmbH. Weinheim, Germany.

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods*, *45*(4), 1234–1245. https://doi.org/10.3758/s13428-013-0324-3

Lithari, C., Frantzidis, C. A., Papadelis, C., Vivas, A. B., Klados, M. A., Kourtidou-Papadeli, C., Pappas, C., Ioannides, A. A., & Bamidis, P. D. (2010). Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. *Brain Topography*, *23*(1), 27–40. https://doi.org/10.1007/s10548-009-0130-5

Livingstone, S. R., & Russo, F. A. (2018). The Ryerson audio-visual database of emotional speech and song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. *PLOS ONE*, *13*(5), e0196391. https://doi.org/10.1371/journal.pone.0196391

Martin, O., Kotsia, I., Macq, B., & Pitas, I. (2006). The eNTERFACE'05 audio-visual emotion database. *Proceedings of the 22nd International Conference on Data Engineering Workshops (ICDEW)*, 8. https://doi.org/10.1109/ICDEW.2006.145

Nayak, J., Naik, B., & Behera, H. S. (2015). Fuzzy c-means (FCM) clustering algorithm: A decade review from 2000 to 2014. In L. C. Jain, H. S. Behera, J. K. Mandal, & D. P. Mohapatra (Eds.), *Computational intelligence in data mining* (pp. 133–149, Vol. 32). Springer India. https://doi.org/10.1007/978-81-322-2208-8_14

Pell, M. D., Jaywant, A., Monetta, L., & Kotz, S. A. (2011). Emotional speech processing: Disentangling the effects of prosody and semantic cues. *Cognition and Emotion*, *25*(5), 834–853. https://doi.org/10.1080/02699931.2010.516915

Perepelkina, O., Kazimirova, E., & Konstantinova, M. (2018). Ramas: Russian multimodal corpus of dyadic interaction for affective computing. In *Affective computing and intelligent interaction workshops* (pp. 501–510, Vol. 10909). Springer. https://doi.org/10.1007/978-3-319-99579-3_52

Plutchik, R. (2013). Measuring emotions and their derivatives. In R. Plutchik & H. Kellerman (Eds.), *The measurement of emotions* (pp. 1–35). Elsevier Science. https://doi.org/10.1016/B978-0-12-558704-4.50007-4

Posner, J., Russell, J. A., Gerber, A., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., & Peterson, B. S. (2009). The neurophysiological bases of emotion: An fMRI study of the affective circumplex using emotion-denoting words. *Human Brain Mapping*, *30*(3), 883–895. https://doi.org/10.1002/hbm.20553

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734. https://doi.org/10.1017/s0954579405050340

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica*, *51*(4), 195–206. https://doi.org/10.1026/0012-1924.51.4.195

Robert, J. (2011). Pydub [Software; Version 0.25.1; retrieved April 10, 2023].

Rozenkrants, B., & Polich, J. (2008). Affective ERP processing in a visual oddball task: Arousal, valence, and gender. *Clinical Neurophysiology*, *119*(10), 2260–2265. https://doi.org/10.1016/j.clinph.2008.07.213

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Sacharin, V., Schlegel, K., & Scherer, K. R. (2012). *Geneva emotion wheel rating study* (Retrieved April 10, 2023). Swiss Center for Affective Sciences, University of Geneva. Geneva, Switzerland. https://archive-ouverte.unige.ch/unige:97849

Santoro, R., Moerel, M., De Martino, F., Goebel, R., Ugurbil, K., Yacoub, E., & Formisano, E. (2014). Encoding of natural sounds at multiple spectral and temporal resolutions in the human auditory cortex. *PLoS Computational Biology*, *10*, e1003412. https://doi.org/10.1371/journal.pcbi.1003412

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. https://doi.org/10.1177/0539018405058216

Schröder, M. (2003). Experimental study of affect bursts. *Speech Communication*, *40*(1-2), 99–116. https://doi.org/10.1016/S0167-6393(02)00078-X

Schuller, B., Vlasenko, B., Eyben, F., Rigoll, G., & Wendemuth, A. (2009). Acoustic emotion recognition: A benchmark comparison of performances. *Proceedings of the 2009 IEEE Workshop on Automatic Speech Recognition & Understanding (ASRU)*, 552–557. https://doi.org/10.1109/ASRU.2009.5372886

Statista Research Department. (2022). *Anzahl deutschsprachiger Menschen* (Retrieved April 10, 2023). https://de.statista.com/statistik/daten/studie/1119851/umfrage/deutschsprachige-menschen-weltweit

Vallat, R. (2018). Pingouin: Statistics in Python [Retrieved April 10, 2023]. *Journal of Open Source Software*, *3*, 1026. https://doi.org/10.21105/joss.01026

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, *26*(4), 557–580. https://doi.org/10.1002/(SICI)1099-0992(199607)26:4⟨557::AID-EJSP769⟩3.0.CO;2-4

Whooley, M. A., Avins, A. L., Miranda, J., & Browner, W. S. (1997). Case-finding instruments for depression. Two questions are as good as many. *Journal of General Internal Medicine*, *12*(7), 439–445. https://doi.org/10.1046/j.1525-1497.1997.00076.x

Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, *38*, 1559–1579. https://doi.org/10.1016/j.pragma.2005.04.012

World Health Organization. (1998). *World Health Organization Five Well-Being Index (WHO-5)* (Manual; retrieved April 10, 2023). Psychiatric Research Unit, WHO Collaborating Centre in Mental Health, and Frederiksborg General Hospital. Hillerød, Denmark. https://doi.org/10.1037/t36093-000

Yang, W., Makita, K., Nakao, T., Kanayama, N., Machizawa, M. G., Sasaoka, T., Sugata, A., Kobayashi, R., Hiramoto, R., Yamawaki, S., Iwanaga, M., & Miyatani, M. (2018). Affective auditory stimulus database: An expanded version of the international affective digitized sounds (IADS-E). *Behavior Research Methods*, *50*(4), 1415–1429. https://doi.org/10.3758/s13428-018-1027-6

4 Study III – Oscillatory and Gaze Signatures of Emotional Speech Processing, Visuo-Spatial Cognition, and Their Interaction in a Near-Realistic Dual-Task MEG Study

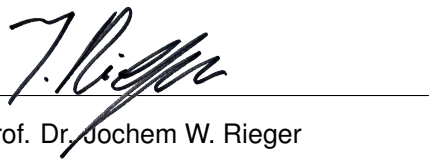
Katharina Lingelbach[1,2], Christoph S. Herrmann[1],
and Jochem W. Rieger[1]

[1] Department of Psychology, University of Oldenburg, Oldenburg, Germany
[2] Fraunhofer Institute for Industrial Engineering IAO, Nobelstraße 12, 70569 Stuttgart, Germany

This article has been published as:
Lingelbach, K., Herrmann, C. S., & Rieger, J. W. (2026). Oscillatory and gaze signatures of socio-emotional speech processing, visuo-spatial cognition, and their interaction in a near-realistic dual-task MEG study. *Imaging Neuroscience, 4*, 1–27. https://doi.org/10.1162/IMAG.a.1134

Author contributions:
Katharina Lingelbach: Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualisation, Writing – original draft, Writing – review & editing. **Christoph S. Herrmann**: Supervision, Validation, Writing – review & editing. **Jochem W. Rieger**: Conceptualisation, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.

The original publication and this chapter are identical in content.
I hereby confirm that Katharina Lingelbach contributed to the aforementioned study as stated above.



Prof. Dr. Jochem W. Rieger
Oldenburg, 30.01.2026

4.1 Abstract

In everyday settings, information processing depends on attentional focus, available cognitive resources, and stimulus characteristics such as its valence. Despite insights from laboratory studies, multisensory information processing in complex, realistic environments remains poorly understood. Using whole-head magnetoencephalography and eye-tracking in a driving simulation, we addressed two research questions: (1) What are the distinct signatures of emotional speech processing and visuo-spatial cognition in the dual task of driving and listening?, and (2) is information processing co-modulated by valence of speech and workload? We hypothesised two antagonistic processing modes: A top-down guided mode for emotional speech processing, and a bottom-up stimulus-driven mode for visuo-spatial cognition under high workload. A third mode was proposed to regulate emotional interference during driving under high visuo-spatial load and limited cognitive capacity. Its involvement is hypothesised to be co-modulated by valence and workload. Spatial clustering of oscillatory source activity supported the hypotheses: During emotional compared to neutral speech, parietal beta-band power increased, indicating predictive listening. In this top-down guided mode, driving-related activity decreased, as indicated by greater parietal alpha- and reduced gamma-band power. During high visuo-spatial workload, linked to the bottom-up stimulus-driven mode, gamma-band power increased in motor and orbitofrontal cortices, whereas beta-band power decreased in motor and temporo-parietal regions. Moreover, increased pupil dilation and decreased gaze dispersion were associated with this bottom-up, stimulus-driven mode and visuo-spatial cognitive demands. Multivariate pattern analyses identified the third regulatory mode reflected in fronto-temporal gamma-band oscillations. It was co-modulated by valence and workload: Under low workload drives, gamma-band activity increased for negative compared to positive speech, indicating inhibitory control. Under high workload drives with heightened visuo-spatial demands, gamma-band activity increased for positive speech, but decreased for negative speech, suggesting reduced cognitive resources and impaired control. To conclude, we found converging evidence of distinct signatures for emotional speech, visuo-spatial cognition, and interference regulation in naturalistic multisensory environments. We propose that the top-down guided mode represents anticipatory listening and socio-emotional cognition, whereas the bottom-up stimulus-driven mode reflects the allocation of resources to driving-related spatial and sensorimotor processing but also cognitive strain under heightened task demands. Fronto-temporal gamma oscillations likely enable flexible up- and down-regulation of emotional speech processing in response to potential interference in complex, naturalistic environments.

Keywords: magnetoencephalography (MEG), neural oscillation, pupillometry, visuo-spatial workload, emotional speech processing, simulated driving, multivariate pattern analysis

4.2 Introduction

Our daily lives require us to constantly attend to, select, and process relevant information from multiple sensory streams, such as listening to an audiobook while monitoring the visual environment and navigating through it during driving.

4.2.1 Top-Down Guided and Bottom-Up Stimulus-Driven Processing Modes Through Oscillations

Oscillatory brain activity plays a crucial role in coordinating neural information processing during cognitive processes (Engel et al., 2001; Fries, 2005), such as attention (Clayton et al., 2015) and working memory (Engel & Fries, 2010; X.-J. Wang, 2010), as well as emotional processing (Briesemeister et al., 2013). Oscillatory modulations in different frequency bands appear to be linked to distinct processing modes; for instance, whether the information is processed externally in a stimulus-driven bottom-up mode or internally in a top-down mode (Engel & Fries, 2010; Engel et al., 2001; Fries, 2005). Gamma oscillations ($>$ 30 Hz) are proposed to be involved in bottom-up sensory processing (Bastos et al., 2015; Ou & Law, 2019), selective attention to relevant stimuli (Jensen et al., 2007; see Herrmann et al., 2010, for review), and regulatory inhibition (Kang et al., 2014; Popov et al., 2012). Specifically, gamma-band activity in the orbitofrontal cortex (OFC) appears to mediate top-down regulatory control and emotion response inhibition by communicating the context-dependent value of a stimulus and decision to the medial prefrontal cortex (mPFC; Kang et al., 2014; Popov et al., 2012; Quirk & Beer, 2006). Oscillatory modulations in the alpha frequency band (8 – 12 Hz) regulate sensory processing by an inhibition-based gating mechanism of stimulus-induced gamma oscillations and cortical engagement (Herring et al., 2019; Klimesch et al., 2007). Beta oscillations have been linked to top-down information processing in the auditory cortex (Hovsepyan et al., 2023; Ou & Law, 2019). Moreover, Engel and Fries (2010) proposed that beta oscillations serve as a mechanism for maintaining the ”status quo”: increased beta oscillations facilitate and stabilise ongoing cognitive and motor processes, whereas unpredicted stimuli reduce beta-band oscillatory power (see also van Pelt et al., 2016).

4.2.2 Multisensory Information Processing in Naturalistic Environments

However, few neurophysiological studies have investigated information processing modes in response to complex, ecologically valid stimuli (Lingelbach et al., 2024; Lingelbach et al., 2023). As a result, the extent to which insights from strictly controlled artificial stimuli and paradigms generalise to real-world situations remains unclear. Notably, neurophysiological responses to dynamic, naturalistic emotional stimuli seem to differ from those evoked by static stimuli (Goldberg et al., 2014; Hamilton & Huth, 2020; Romeo et al., 2022).

The concurrent task of driving while listening to speech represents a safety-critical naturalistic scenario of multisensory information processing and frequent attentional shifts (see Haghani et al., 2021; Palmiero et al., 2019, for reviews). This dual-task scenario has been shown to reduce activity in brain regions associated with driving and decrease driving performance while simultaneously recruiting areas involved in speech processing

(Fort et al., 2010; Sasai et al., 2016; Sonnleitner et al., 2012). Parietal regions linked to spatial processing and cognition, along with bilateral occipital areas, are associated with driving (Navarro et al., 2018; Sasai et al., 2016; Spiers & Maguire, 2007), whereas temporal regions and the inferior frontal gyrus are associated with speech processing (Friederici, 2012).

In their behavioural study, Nowosielski et al. (2018) observed improved reaction times to hazards when participants listened to audiobooks during easy drives, but not during challenging ones. This finding underscores the need to examine how secondary listening tasks interact with different levels of driving-induced visuo-spatial workload.

Emotional Speech Processing During Driving

Speech conveys meaning, intention, and emotion through both semantic and paralinguistic components. The semantic component refers to *what* is said, while the paralinguistic component relates to *how* it is expressed. The latter encompasses prosody, that is, the melodic aspect of speech (pitch, duration, and intensity; Belyk & Brown, 2014).

Emotional speech processing is proposed to occur via two neural pathways (Phillips et al., 2003; Schirmer & Kotz, 2006; Tucker et al., 1995; Wildgruber et al., 2006): One pathway operates automatically, primarily involving the superior temporal gyrus (STG) and sulcus (STS). The other relies on deliberate, controlled evaluation, engaging regions such as the right inferior frontal gyrus (IFG), OFC, dorsolateral prefrontal cortex (dlPFC), and parietal areas (see Belyk & Brown, 2014; and Witteman et al., 2012, for meta-analyses; see Grandjean, 2021, for review).

Although the impact of conversations on driving has been widely studied at the behavioural level, and to a lesser extent, electrophysiologically (Haghani et al., 2021; Palmiero et al., 2019), very little research has investigated how emotional speech with positive and negative valence impacts attention and neural information processing during driving. Assessing gaze behaviour during driving with fear-related conversations, Briggs et al. (2011) reported decreased driving performance, fewer fixations, and the phenomenon of visual tunnelling, typically reported for increased cognitive workload during driving (Nunes & Recarte, 2002).

4.2.3 Co-Modulation Effects of Emotional Speech and Visuo-Spatial Workload

A growing body of research is investigating how cognitive and emotional processes interact when cognitive resources are limited (Brockhoff et al., 2022; Cromheeke & Mueller, 2014; Dolcos et al., 2011; Schweizer et al., 2019).

Previous studies indicate that the availability of cognitive resources (i.e., attentional control and the capacity to process information; Baddeley, 1992; Cowan, 2017; Wickens, 2014) influences how emotional stimuli are processed (Brockhoff et al., 2022; Dong et al., 2024; Tavares et al., 2016). Dual-task studies indicate that emotional processing, and thus valence-specific processing, only occurs when cognitive workload is low (Dong et al., 2024; Lingelbach & Rieger, 2025; Tavares et al., 2016). When cognitive workload is high and cognitive resources are scarce, attentional control promotes goal-directed behaviour by directing perception,

processing, and response selection (Cromheeke & Mueller, 2014; E. K. Miller & Cohen, 2001). In these circumstances, the processing of task-irrelevant emotional content appears to be reduced (Brockhoff et al., 2022; Dong et al., 2024; Lingelbach & Rieger, 2025; Tavares et al., 2016). This phenomenon was evident even in cross-modal studies employing a visual task alongside auditory emotional stimuli (Mothes-Lasch et al., 2012). Mothes-Lasch et al. (2012) observed valence-specific processing for negative words during low, but not high, perceptual load in an event-related study. However, a functional near-infrared spectroscopy (fNIRS) study by Lingelbach et al. (2023) showed that participants attempted to inhibit naturalistic negative speech irrespective of the current workload level, when presented continuously in a block design. Moreover, the study reported reduced brain activity in the left temporal pole with increasing workload, suggesting a decline in the efficiency of inhibitory control (Lingelbach et al., 2023).

Several prefrontal areas play a key role in goal-directed attentional control to regulate the effects of distracting emotional information. These areas comprise the dlPFC, ventrolateral (vlPFC), and mPFC, anterior cingulate cortex (ACC), OFC, temporal pole, and IFG (Kang et al., 2014; Lingelbach & Rieger, 2025; Lingelbach et al., 2023; Ochsner et al., 2012; Popov et al., 2012).

Gaze Behaviour During Multisensory Information Processing

Gaze behaviour is widely studied as an indicator of information processing and cognitive capacity in multisensory environments of varying complexity.

Faure et al. (2016) examined the relationship between blink activity and dual-tasking during driving. They found that blink frequency declined with increasing driving complexity. However, a secondary, auditorily presented task increased blink frequency. To date, no study has investigated the effects of emotional speech or the co-modulation of emotional speech and visuo-spatial complexity on blink behaviour.

Two pupillary indices, the change of pupil diameter (i.e., pupil dilation) and discontinuities in its rate of change (so-called Index of Cognitive or Pupillary Activity; hereafter IPA; see Duchowski, 2018), appear to be tied to different processing modes during dual-tasking (Vogels et al., 2018). Vogels et al. (2018) found that discontinuities in pupil size changes decreased when cognitive resources are distributed across multiple tasks (i.e., in dual tasks), whereas pupil dilation increased. While the latter finding aligns with the well-established relationship between cognitive workload and pupil dilation (Beatty, 1982), the former is unexpected and contradicts previous results in pupillometry (Czerniak et al., 2021; Duchowski, 2018). Surprisingly, discontinuities increased significantly with linguistic complexity but not driving complexity (Vogels et al., 2018; see also Demberg & Sayeed, 2016).

While the relationship between pupil dilation and neural markers of attention and visuo-spatial processing is well established (Alnæs et al., 2014; Joshi et al., 2016; Vinck et al., 2015), the link between pupillary discontinuities and neural activity remains unexplored. Furthermore, their role in attention and information processing is still unclear in dual-task contexts (Demberg & Sayeed, 2016; Vogels et al., 2018).

4.2.4 Research Question and Hypotheses

To date, no study has integrated neural oscillatory and pupillary change discontinuities to investigate attentional allocation and information processing modes. Furthermore, information processing during emotional speech and visuo-spatial workload under ecologically valid conditions, as well as potential interaction effects between these factors, remain unexplored, despite their relevance for driving safety. In their Road Safety Policy Framework, the European Commission identified distracted driving as a leading cause of road crashes, surpassing speed and alcohol. They emphasised the need for further research to understand and identify safety-critical mental states to develop effective warning technology (European Commission & Directorate General for Mobility and Transport, 2020). This strongly motivates the present foundational magnetoencephalography (MEG) study. Insights into the neurophysiological interplay of cognitive workload, resource capacity, and inference regulation in individual drivers can inform the design of personalised assistive systems that support safe and effective driving (Chavarriaga et al., 2018; Scheunemann et al., 2019).

Thus, this study combined whole-head MEG, offering high temporal and fair spatial resolution, with eye-tracking to address two outstanding questions:

1. R.1. What are the distinct signatures of emotional speech processing and visuo-spatial cognition in the dual task of driving and listening?

2. R.2. Do the valence of emotional speech and workload associated with the visuo-spatial cognition co-modulate information processing?

We hypothesised two distinct, antagonistic information processing modes:

- H.1. Emotional (low, LV, and high, HV, valence) compared to neutral speech (neutral valence, NV) is expected to increase oscillatory beta-band power in parietal regions (predictive listening; Hovsepyan et al., 2023; Ou & Law, 2019), and the number of discontinuities in pupil size changes (IPA; Vogels et al., 2018). Concurrently, oscillatory gamma-band power is expected to decrease in parietal areas linked to spatial processing (Fort et al., 2010; Sonnleitner et al., 2012).

- H.2. High compared to low visuo-spatial workload is expected to increase gamma- and decrease alpha-band power in occipital, parietal, and motor areas as a response to greater task demands. We hypothesise that the increased workload detrimentally affects driving performance. Concurrently, parietal beta-band power associated with predictive speech tracking and internalised processing is expected to decrease (Engel & Fries, 2010; Ou & Law, 2019; Sakihara et al., 2014; Sonnleitner et al., 2012). For gaze behaviour, we hypothesised increased pupil dilation, as well as decreased blink activity (Faure et al., 2016), and distribution of fixations (i.e., visual tunnelling; Briggs et al., 2011).

Regarding research question 2, we expect an interaction between the valence of emotional speech and workload (Brockhoff et al., 2022; Lingelbach & Rieger, 2025; Schweizer et al., 2019):

- H.3. Valence-specific processing and appraisal of positive speech should occur only under low workload, when cognitive resources are available. Under low workload, positive compared with negative speech is assumed to lead to higher ratings in valence and lower ratings in arousal, frustration, effort, distraction, and IPA scores. Under high workload, frontal control mechanisms in the form of gamma oscillatory activity in the mPFC and OFC are expected to regulate the processing and suppress appraisal of positive speech (Kang et al., 2014; Ochsner et al., 2012; Popov et al., 2012; Quirk & Beer, 2006). This response resembles the regulation of negative speech, which is assumed to occur irrespective of workload, to reduce emotional interference (Lingelbach et al., 2023).

4.3 Methods

4.3.1 Participants

A total of 48 volunteers participated in the study ($M_{age} = 25.25$, $SD = 4.01$, range = 19 – 38 years, 25 females and 23 males). Before the study, participants were screened to ensure they met the pre-defined inclusion criteria. The criteria included having a valid driver's license, normal or corrected-to-normal vision, no neurological or psychiatric disorders, no history of psychoactive substance use in the last 14 days, magnetic resonance imaging (MRI) and MEG compatibility, and being a native German speaker. Participants had on average 7.9 years of driving experience ($SD = 3.9$, range = 1 – 20 years). They provided written informed consent before the experiment and received monetary compensation, with a bonus for sound driving behaviour and timely arrival at their destination. The experiment was approved by the Commission for Research Impact Assessment and Ethics at the University of Oldenburg, Germany (Ref: EK/2018/070), and conducted following the Declaration of Helsinki. For the MEG source space analysis, we excluded seven participants ($N = 41$, $M_{age} = 25.44$, $SD = 4.20$, range = 19 – 38 years, 20 females and 21 males) due to a missing individual MRI scan caused by a panic attack in the scanner ($n = 1$), and excessive movements during driving ($n = 6$; number of epochs per condition left after movement artifact rejection < 30). For the eye-tracking analysis, five participants were excluded due to difficulties in obtaining a stable eye-tracking signal ($N = 43$, $M_{age} = 25.35$, $SD = 3.97$, range: 19 to 38 years, 24 females and 19 males).

4.3.2 Procedure and Material

The experiment was controlled using Python 3.7. The driving scenarios were simulated with the software SILAB 6.5 (Krueger et al., 2005). The software communicated with a local Python TCP/IP socket using a binary protocol to transmit experimental triggers. We used a parallel port and Expyriment (version 0.10.0) to send synchronisation and response triggers. The triggers were stored with the MEG recordings.

Driving Simulation

The driving stimulation task was displayed using rear projection on a screen measuring 750 x 428 mm inside the magnetically shielded room of the MEG with a resolution of 1920 x 1080 pixels (PROPixx DLP LED

Projector, 480 Hz refresh rate and VPixx controller). The back-projection screen was positioned 115 cm away from the participant's eyes. The screen covered 15.07° of visual angle in width and 8.73° in height from the centre. The eye-tracker was positioned 103 cm away from the participant's eyes at an angle of 22.5°.

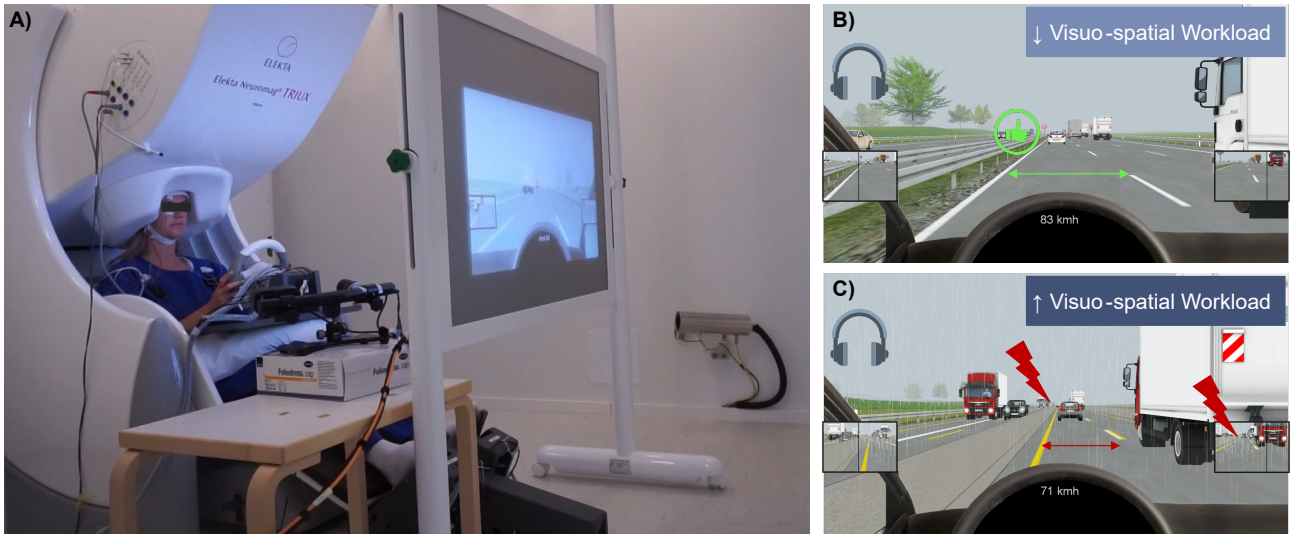
At the beginning of the experiment, participants underwent a three-minute resting-state recording with their eyes open and fixated on a fixation cross presented centrally on a grey background. They were then given 10 minutes to familiarise themselves with the driving simulator and equipment in a practice drive.

The participants controlled the ego car in the driving simulation via a standard interface consisting of a vehicle mock-up with a throttle, brake pedal, and steering wheel (Figure 4.1A).

The experimental task involved six driving blocks of 12 km each, plus a 0.75 km baseline at the beginning and a 0.75 km transfer section at the end of each block. The speed limit of the baseline and transfer sections was set at 50 km/h. In other sections, the speed limit changed seven to eight times per block (mean: 70 km/h; range: 50 – 90 km/h). Radar controls, indicated by a red flash, and speed warnings ensured that participants did not drive too fast. Speed warnings were displayed on the windscreen whenever participants exceeded $10\ \%$ of the speed limit for more than 5 seconds. Collisions with other cars triggered a yellow flash as feedback. Participants were instructed to drive as fast as possible to their destination without violating traffic rules. For violations (radar flash or collision), a penalty was deducted from the monetary bonus.

Figure 4.1

Experimental Set-Up of the Driving Simulator and Standard Vehicle Mock-Up With a Steering Wheel, Throttle, and Brake Pedal



Note. Participants were seated in the MEG inside the magnetically shielded chamber.

Each block represented one of the six conditions in a 2 × 3 factorial design. The two factors were: (1) *visuo-spatial workload* during driving (low vs. high); and (2) *emotional auditory speech* from one or two

speakers with low, high, or neutral valence. Workload was manipulated via visuo-spatial attention and driving difficulty. Low visuo-spatial workload (LW) sequences featured no construction areas (lane width = 3.75 m), three lanes, slightly foggy, dry weather, and predictable other car agents (Figure 4.1B). High visuo-spatial workload (HW) sequences included construction areas with two narrower lanes (lane width = 2.75 m), slightly foggy, rainy weather, and less predictable agents exhibiting risky driving behaviours (e.g., sudden braking and tailgating; Figure 4.1C). Traffic density was comparable across conditions, with six to eight vehicles randomly distributed within a 1000 m radius of the ego car (similar to Scheunemann et al., 2019). The ego car was a typical passenger car, measuring 4.2 m in length and 1.8 m in width.

Visuo-spatial workload conditions alternated across blocks and were counterbalanced across participants. Valence of speech was randomised, ensuring no consecutive repetitions of the same valence. A representative video for each condition is provided in the Open Science Framework (OSF) project repository (https://osf.io/um6vw/).

Auditory Stimulation

Emotional naturalistic speech sequences were taken from the validated GAUDIE database (German AUDItory Emotional Database; Lingelbach et al., 2024). We included 11 stimuli from each condition in the study (metrics of the stimulus set can be found in the Supplementary Material 8.1). Negative speech stimuli comprised television excerpts featuring arguments between two characters, positive stimuli consisted of excerpts from comedian shows, while neutral stimuli comprised weather forecasts. The audio sequences were digital-to-analogue converted at a sampling frequency of 44.1 kHz using the external soundcard Fireface UCX (RME) along with an HB7 headphone amplifier (Tucker-Davis Technologies Inc., Progress Corporate Park, Alachua, Florida). The output was dichotically presented via ER3-14A insert earphones (Etymotic Research Inc.), equipped with single-use foam tips.

Questionnaires and Ratings

At the end of each driving block, participants rated their perceived effort and frustration using a modified subscale from the NASA Task Load Index (TLX, with a continuous scale ranging from 0 to 20; Hart & Staveland, 1988). They also assessed their overall valence and arousal, as well as the perceived valence and arousal of the speech, using the Self-Assessment Manikin (SAM) subscales (continuous scale ranging from 0 to 1; modified from Bradley & Lang, 1994). Additionally, participants estimated the perceived amount of distraction from the auditory speech (continuous scale ranging from 0 to 20), and answered control questions regarding the conversational content of the speech (answer option: *True*, *False*, or *Not Perceived or Memorised*).

4.3.3 Data Acquisition

Magnetoencephalogram

Neuromagnetic signals were recorded using a 306-channel whole-head MEG system (Elekta Neuromag Triux, Elekta Oy, Helsinki, Finland) with 102 magnetometers and 204 orthogonal planar gradiometers. The MEG

system was housed in a magnetically shielded chamber (Vacuumschmelze, Hanau, Germany). The dewar was positioned at 68°, and participants sat upright beneath the MEG sensors. Five head position indicator (HPI) coils were attached to the participants' heads for continuous head position tracking. To co-register the MEG signals with the structural T1 MRI scans, coil positions and anatomical landmarks (fiducials: nasion, left pre-auricular (LPA), and right pre-auricular (RPA) points) were digitised, along with at least 200 head-shape samples (Whalen et al., 2008), using the Polhemus Fastrak® (Polhemus, Colchester, VT, USA). The MEG signals were recorded without internal active shielding, at a sampling rate of 1 kHz, and with an online band-pass filtering between 0.1 Hz and 330 Hz.

Magnetic Resonance Imaging (MRI) Structural Scans

Two structural T1-weighted MRI scans were obtained from each participant using a Siemens Magnetom Prisma 3.0 Tesla MRI machine (Siemens, Erlangen, Germany) with a 3D T1-weighted sequence (MPRAGE, TR = 2000 ms, TE = 2.07 ms, flip angle = 9°, voxel size = 0.75 x 0.75 x 0.75 mm^3, GRAPPA = 2, field of view = 240 x 240 mm, 224 sagittal slices, fat-saturated, TA = 7:45 min). The two T1 images were averaged to improve the signal-to-noise ratio (SNR), segmented into specific brain tissues, and individual brain surfaces were reconstructed for source localisation using FreeSurfer (version 6.0.0; Dale et al., 1999; Fischl et al., 1999).

Eye-Tracking

Gaze behaviour was recorded during the experiment using the infrared remote eye-tracking device EyeLink 1000 Plus (SR Research Ltd., Ottawa, Canada) with a sampling rate of 1 kHz. At the beginning of each experiment, the eye-tracker was calibrated using a 9-point calibration. At the beginning of each block, a drift correction was performed as implemented in the EyeLink software. We kept the average calibration error below 0.5° of visual angle, with a maximum calibration error of 1.0°, and drift corrections between driving blocks below 5.0°. In cases of larger deviations, the eye-tracker was recalibrated.

Driving Behaviour

Behavioural data from the simulator included the ego car's vehicle position coordinates (X, Y, Z in m), speed (m/s), acceleration (m/s^2), velocity (m/s), steering wheel angle (radians/s), and the positions of the acceleration and brake pedals (continuous scale: not pressed: 0 to fully pressed: 1). The recordings were conducted via SILAB.

4.3.4 Data Analysis

We analysed the data using Python 3.9 and MNE-Python (version 1.6.1; Gramfort et al., 2013).

MEG Preprocessing in Sensor Space

For preprocessing, the neuromagnetic signal was first decomposed into spatiotemporal components of internal and external origin using the Maxwell filter (MNE-Python default settings; Taulu & Simola, 2006; Taulu et al., 2005). In this step, flat and noisy channels were identified and reconstructed, and the data were aligned to a unified coordinate system across driving blocks. Additionally, head movements were corrected by realigning the signals to the initial head position using continuous HPI coil tracking. We examined the head movement correction by calculating the HPI coil goodness of fit (scale: 0 – 1), the average 3D deviation from the initial head position, and deviations between consecutive samples. The mean HPI goodness of fit exceeded 0.99. Average deviation from the initial position was 1.76 mm (SD = 0.86 mm), with $67.05\ \%$ of values below 2 mm. The mean deviation between consecutive samples was 0.13 mm (SD = 0.07 mm), with $96.67\ \%$ below 0.5 mm. In the next preprocessing step, raw data were downsampled to 100 Hz and band-pass filtered using a 4^{th}-order infinite impulse response (IIR) Butterworth filter with a low cut-off frequency set at 0.1 Hz and a high cut-off at 42 Hz. Next, data from the driving blocks were concatenated. We performed a semi-automated independent component analysis on the concatenated raw data using MNE-Python (Gramfort et al., 2013), and the extended infomax algorithm (T.-W. Lee et al., 1999). Electrocardiac signals were reconstructed from the MEG signals, and the blink-related artefacts were identified using information from measured electrooculography (i.e., additional passive Ag–AgCl electrodes recording horizontal and vertical eye movements). Further contaminated components (e.g., muscle and non-physiological artefacts) were manually selected based on visual inspection of the topography, time course, and power spectral intensity (Chaumon et al., 2015; Hipp & Siegel, 2013). The number of independent components was determined to retain at least $99\ \%$ of the explained variance ($M = 74.21$, $SD = 2.12$, range = 68 – 78). On average, 8.88 components (SD = 1.72; range = 5 – 13) were removed per participant before back-projecting the signals into sensor space. After preprocessing, the continuous signals were segmented into baseline sections of 0.75 km and experimental condition sections of 12 km, based on the trigger signal marking the start of each block. From the segments, we created epochs of non-overlapping windows with an amplitude rejection for noisy epochs exceeding a threshold of 4000 fT for magnetometers and fT/cm for gradiometers. The number of epochs was equalised across conditions by employing a method that minimised timing discrepancies across trial lists, ensuring an identical trial count per condition.

MEG Signal Reconstruction to Source Space

To model the head's geometry, a boundary element model (BEM; Mosher et al., 1999) was constructed using the inner skull surface extracted via the FreeSurfer watershed tessellation algorithm (Ségonne et al., 2004). A single-shell mesh was generated from this surface and assigned a conductivity of 0.3 S/m. The source space was created as a uniformly distributed grid of dipoles across the cortical surfaces according to the MNI305 (Montreal Neurological Institute) space (Collins et al., 1994), as implemented in FreeSurfer. We used icosahedron subdivisions of 5 mm, resulting in approximately 10242 sources per hemisphere and a source spacing of 3.1 mm (i.e., a surface area per source of 9.8 mm^2). We co-registered the MEG data

with the MRI anatomy derived from individual T1 scans using a semi-automated approach following Houck and Claus (2020). Firstly, fiducials (i.e., LPA, RPA, and nasion) were estimated in the MRI head space and translated into the participant's MRI coordinate space. The estimated landmarks derived from a template brain (Dale et al., 1999; Gramfort et al., 2013) were visually inspected and manually subject-wise adjusted. Next, we fitted the fiducials from the MRI and MEG coordinate space by scaling, translation, and rotation (relative weight for nasion = 2; LPA and RPA = 1). To align the head shape points and MRI, the iterative closest point (ICP) algorithm was applied (20 iterations; overall weight = 1). Afterwards, head shape points with a distance larger than 5 mm to the MRI skin surface were omitted, and the ICP algorithm was repeated (20 iterations; naison weight = 10; other points' weight = 1). The co-registration resulted in a mean distance between head shape points and MRI skin surface of $M = 1.59$ ($SD = 0.28$) mm across participants. The forward model was computed by estimating the lead-field matrix using the coregistration model, source space, and BEM solution for each participant (Hämäläinen et al., 1993). The noise-covariance matrix was estimated from a combined empty room measurement of three minutes in duration, performed before and following the experiment. The empty room data were preprocessed similarly to the experimental MEG data. The regularisation was applied with the Ledoit-Wolf shrinkage, wherein the parameter alpha was optimised through a cross-validated search. The rank was previously calculated from the normalised data (Gramfort et al., 2013).

Mass-Univariate Permutation-Based Spatial Clustering and Multivariate Pattern Analysis

We investigated the two research questions with two approaches: To identify neural information processing modes during emotional speech and visuo-spatial cognition (R.1; main effects), we investigated differences in modulated oscillatory power in the alpha, beta, and gamma frequency bands with dynamic imaging of coherent sources (DICS) and permutation-based spatial clustering in source space. For the second research question (R.2; interaction effects), we expected rather small effect sizes at the brain level due to individual regulation strategies (Morawetz & Basten, 2024), as well as anatomical and functional variability in activation patterns across subjects (Michalke et al., 2023). Thus, we approached R.2 by applying multivariate pattern analysis (MVPA) using common spatial pattern (CSP) combined with a linear discriminant analysis (LDA). For converging evidence, we also employed mass-univariate permutation-based spatial clustering in a supplementary analysis.

Permutation-based spatial clustering is a data-driven method used to identify significant clusters of vertices that differ in their response across experimental conditions, while controlling for multiple comparisons (Maris & Oostenveld, 2007). To detect meaningful differences between conditions, a significance threshold is defined (here: $p < .05$), and measuring positions (here: vertices) are grouped based on spatial adjacency. Condition labels are then randomly reassigned across clusters in permutations (here $n = 5,000$) to generate a randomised dataset. For our 2 $\times$ 3 factorial design, an *F*-value is computed for each permutation using a repeated-measures analysis of variance (rmANOVA; one-sided). A cluster is considered significant if the sum of *F*-values in the original data exceeds the 95$^{\text{th}}$ percentile ($p < .05$) of the *F*-value distribution from the randomised data (Maris & Oostenveld, 2007).

MVPA is a data-driven method that employs machine learning (ML) to integrate multiple measurement

sources within a multidimensional framework while taking into account inter-individual neural variability (Haufe et al., 2014; Holdgraf et al., 2017; Marsicano et al., 2024). When using linear supervised ML models combined with inverse mapping techniques (Blankertz et al., 2008; Grosse-Wentrup & Buss, 2008; Haufe et al., 2014), decoding patterns that differentiate between conditions in the classification can be derived and physiologically interpreted (Holdgraf et al., 2017). These patterns can subsequently be localised in source space (Gramfort et al., 2013).

Modulated Oscillatory Power with Dynamic Imaging of Coherent Sources (R.1). To localise modulated oscillatory power of emotional speech processing and visuo-spatial cognition (R.1), we used the beamformer DICS (J. Gross et al., 2001; van Vliet et al., 2018). This frequency-domain beamformer constructs spatial filters in source space to estimate frequency-specific activity from a target location, while suppressing contributions from other sources (van Veen et al., 1997; van Vliet et al., 2018). For DICS, cross-spectral density (CSD) matrices were computed using Morlet wavelet decomposition on 5-second epochs (3 cycles, frequencies from 1 – 42 Hz). The number of epochs per condition was on average $M = 99.83$ ($SD = 16.01$, range: 33 – 115; $M_{drop} = 19.15\ \%$). The CSDs and the forward model were used to construct the spatial filters for each cortical grid point. These spatial filters were then applied condition-wise to calculate the relative change in each condition from its baseline across frequencies (van Vliet et al., 2018). To combine magnetometer and gradiometer signals, data were spatially pre-whitened using a noise-estimating covariance matrix from the empty room measurement (Engemann & Gramfort, 2015; Gramfort et al., 2013).

Oscillatory activity within canonical frequency bands is superimposed upon a broadband, non-oscillatory (aperiodic) background component (Donoghue et al., 2020; K. J. Miller et al., 2009). This component exhibits a scale-free, 1/*f*-like spectral profile (He et al., 2010; K. J. Miller et al., 2009). It is assumed to be evoked by the asynchronous summation of postsynaptic potentials across large neuronal populations, reflecting the underlying balance of excitatory and inhibitory synaptic activity and their respective time constants (e.g., He et al., 2010; K. J. Miller et al., 2009). In recent years, increasing attention has been directed towards the aperiodic aspects of neural power spectra (Donoghue et al., 2020; Gerster et al., 2022; Lu et al., 2024; Thuwal et al., 2021). Importantly, the 1/*f* signal appears to be also modulated by task-related experimental manipulations and may reflect not only background neural noise but also physiologically meaningful activity (Donoghue et al., 2020; Lu et al., 2024; C. Zhang et al., 2023). Consequently, changes in the aperiodic components can confound or even obscure modulations in oscillatory power above the 1/*f*-like activity (Donoghue et al., 2020; Gerster et al., 2022). This is particularly true for experimental paradigms that are not stimulus-locked relative to a pre-stimulus baseline. These considerations motivate the separation of the aperiodic and oscillatory components above this 1/*f*-like part for independent investigation.

To separate 1/*f*-like activity from the superimposed oscillatory activity, we visually inspected the power spectra to ensure there were no apparent objections to decomposing the signal components as proposed by Gerster et al. (2022). Afterwards, we estimated the aperiodic component, including the offset and slope of the 1/f decay, using the FOOOF algorithm (version 1.1.0; Donoghue et al., 2020). This was performed for each participant, condition, and vertex with the following parameters: peak width limits = 2 – 6 Hz, maximum number

of peaks = 7, absolute peak detection threshold = 0.0, relative peak detection threshold = 1.0, and a fixed mode. The aperiodic component was then subtracted from the power spectral density in linear space (Gyurkovics et al., 2021a).

Supplementary Figure 8.4 shows the power spectrum separated from the $1/f$-like signal (solid lines) and aperiodic component (dotted lines; A – C) for the single effect (A) and main effect conditions (B: visuo-spatial workload; C: emotional speech).

To obtain the power within the frequency bands of interest, individual band peaks were estimated from resting-state data in sensor space. Similar to the processing of the conditions, non-overlapping 5-second epochs were extracted from the resting-state data, and power was calculated using the multitaper method with a bandwidth of 2 and averaged across epochs for both sensor types. Peak frequencies for the alpha, beta, and gamma bands were then identified for each sensor using the FOOOF algorithm (Donoghue et al., 2020), constrained to the following cut-off ranges: 8 – 12 Hz (alpha), 15 – 25 Hz (beta), and 30 – 42 Hz (gamma). Peak frequencies were averaged across all sensor positions and types, and the source-space frequency band power was calculated using a 2 Hz bandwidth centred on the individual peak. The average R^2 of the power parametrisation models was 0.97 ($SD = 0.01$), with a mean error of 0.03 ($SD = 0.01$). Descriptive statistics of the band centre frequencies are provided in Supplementary Table 8.1. The narrowband oscillatory power was averaged for each frequency band, and the spatial distribution of the individual oscillatory power of each condition and frequency band was morphed to the average FreeSurfer brain template fsaverage (Fischl et al., 1999; Gramfort et al., 2013). This morphing transforms the source space of individual subjects into a common source space, enabling group-level statistical analysis. Grand averages of the DICS localised oscillatory source power per frequency band and main effect condition (R.1) are visualised in Supplementary Figure 8.5 (with subtraction of the aperiodic part) and Supplementary Figure 8.6 (without subtraction of the aperiodic part).

For the non-parametric permutation-based clustering statistic, we tested for significant main effects of emotional speech and visuo-spatial workload using an rmANOVA (Maris & Oostenveld, 2007). Potential interaction effects were also examined to rule out cross-over interactions that could compromise the interpretability of the main effects. *F*-values of significant clusters were projected on a 3D brain (*fsaverage*). In case of a significant main effect, oscillatory power was averaged across the vertices of the significant *F*-test cluster, and pairwise Wilcoxon signed-rank tests (non-parametric alternative to a dependent *t*-test; scipy version 1.5.0) were conducted for the respective contrasts, with correction for family-wise error using the false discovery rate (FDR with the Benjamini-Hochberg method; statsmodels version 0.13.5).

Exploratory analyses of changes in the aperiodic components in response to emotional speech and visuo-spatial workload are presented in the Supplementary Material 8.7.

To investigate the relationship between significant neural markers of information processing during emotional speech and visuo-spatial workload (R.1), we conducted correlation analyses using Spearman rank correlations (r_s). Average cluster source band power was computed over the vertices of significant clusters of the main effects of emotional speech and workload. To account for multiple correlations, the significance level was set at $\alpha = 0.01$.

Interaction Effect of Valence and Workload on Modulated Oscillatory Power (R.2). To investigate the second research question (R.2; interaction effects), we selected a shorter epoch length of 2 s (non-overlapping) to increase the number of epochs per condition for the ML decoding. The number of epochs was equalised across conditions with an average of $M = 275.12$ ($SD = 11.54$, range: 237 – 296; $M_{drop} = 8.17\ \%$) per condition. In the MVPA, the cleaned raw data were band-pass-filtered using a linear-phase finite impulse response (FIR) filter with zero-padding in the gamma-band frequency range (cut-offs: 30 and 42 Hz) before epoching. Next, we applied MVPA to the epoched band-pass filtered sensor-space data using a multiclass extension of CSP (Blankertz et al., 2008; Grosse-Wentrup & Buss, 2008, as implemented in MNE-Python), combined with a linear discriminant analysis (LDA; implemented in scikit-learn 1.5.2). The decoding was performed using the LDA with least-squares solution as solver to discriminate the four experimental conditions in a within-subject four-class classification (emotional speech/visuo-spatial workload conditions: HV/LW, LV/LW, HV/HW, LV/HW). We selected the first four CSP components (log-transformed) as input for the LDA. Covariance matrices in both the multiclass CSP extraction and LDA were regularised using the Ledoit-Wolf method. We used the Pipeline method to streamline all preprocessing and classification steps (scikit-learn 1.5.2). Each participant's data (shape: n_{epochs}, 306 channels, 200 time points) was partitioned into epochs for training and testing using repeated stratified ten-fold cross-validation with three repetitions (30 folds total). To evaluate decoding performance, we used the *F*1 score and analysed confusion matrices across participants. The subject-wise mean *F*1 scores and their $95\ \%$ confidence intervals (CIs) were computed by bootstrapping across cross-validation folds using a Monte Carlo simulation (MCS with 5,000 iterations; Cumming & Finch, 2005). The spatial patterns of the CSP component used in the decoding were derived from the fitted linear models and averaged across cross-validation folds. Further methodological information regarding CSP for multiclass decoding can be found in the Supplementary Material 8.4

Although DICS is well-suited for investigating oscillatory power modulations in source space, it estimates spatial filters using a common CSD matrix that combines information across all experimental conditions. However, creating a common CSD matrix is not applicable when transforming CSP components derived from MVPA decoding to source space. Therefore, we chose Minimum Norm Estimate (MNE) as inverse solution for source reconstruction to investigate R.2 and H.3. MNE estimates the most likely distribution of neural sources by selecting the one with the minimum overall current amplitude (as measured by the L2 norm) that explains the observed magnetic fields (Hämäläinen et al., 1993). It is particularly useful when the source configuration is complex or unknown (Hauk, 2004). For the localisation of the CSP component patterns in source space, the SNR was set at 3. To estimate gamma-band power from the inverse solution of the epoched data, we used the multitaper method with discrete prolate spheroidal sequence (DPSS) windows, a 2 Hz bandwidth, and an SNR of 1 (Gramfort et al., 2013). Gamma-band estimates were used to interpret component patterns in relation to the experimental conditions. In both cases (i.e., CSP component patterns and gamma-band oscillatory power), the depth prior was estimated from the data using a weighting exponent of 0.8. We applied a loose orientation constraint with a weighting factor of 0.2 (Lin et al., 2006), and computed the norm of the resulting loose orientations. Decoding patterns and gamma-band power estimates in source space were then aligned across participants by morphing them to the *fsaverage* template brain (Fischl et al., 1999; Gramfort et al.,

2013). Grand averages were computed by weighting each participant's activation pattern according to their test decoding *F*1 score, thereby enhancing the influence of participants with higher classification performance. Subsequently, the resulting averages were rescaled to a 0 – 1 range and visualised on the 3D *fsaverage* brain template.

To examine how the CSP component patterns relate to the emotional speech and visuo-spatial workload conditions, gamma-band source power was averaged across vertices whose absolute decoding values exceeded the 90^{th} percentile in the CSP pattern. Differences between conditions were assessed using Wilcoxon signed-rank tests with FDR correction, and visualised via bootstrapped condition means and their CIs (MCS with 5,000 iterations).

In a final supplementary analysis, we compared the MVPA results with a traditional mass-univariate, permutation-based clustering (Maris & Oostenveld, 2007). Consistent with the approach used for R.1, a rmANOVA with a one-sided omnibus *F*-test was applied to identify clusters reflecting interaction effects in gamma-band source power localised with MNE. In case of a significant interaction effect, oscillatory power was averaged across the vertices of significant clusters, and Wilcoxon signed-rank tests were performed (FDR-corrected). As before, results were visualised using bootstrapped condition means and their CIs (MCS with 5,000 iterations).

Analysis of Subjective Ratings

To investigate our second research question of a co-modulation on the subjective experience (R.2, H.3), we analysed self-reported ratings collected at the end of each driving block. The effort and frustration ratings, originally collected on a 0 – 20 scale in the NASA-TLX, as well as the distraction scale, were rescaled to a 0 – 1 range to match the scale used for the SAM, and to facilitate interpretation. Recall and missed content in the conversation questions were dummy-coded for further supplementary analyses.

Analysis of Driving Behaviour

We used the following measures to assess driving performance: changes in (1) acceleration and (2) steering wheel angle, both computed as the root mean square of successive differences (*RMSSD*); (3) average brake actuation; (4) traffic rule violations, including the number of radar flashes and collisions; and (5) mean lane deviation from the centre line, excluding two seconds around each lane change. These measures were aggregated into an overall driving performance score. This was done by first normalising each measure to a 0 – 1 range, averaging across measures, and inverting the aggregated score so that higher values indicated better performance. The final performance score was then rescaled once more, yielding values ranging from 0 (indicating poor performance) to 1 (indicating high performance).

Analysis of Gaze-Related Measures

To assess gaze-related changes in information processing under varying visuo-spatial workload and emotional speech, we focused on four measures: (1) gaze dispersion, (2) blink rate per second, and two

pupillary measures, (3) the average pupil dilation, and (4) IPA, reflecting the number of abrupt, sharp fluctuations in pupil diameter per second. All measures were baseline-corrected by subtracting the value obtained during the baseline block from the value in the immediately following experimental block.

Blinks, including minimal, missing, or distorted pupil signals due to eyelid closure, were identified by an online event parser of the EyeLink eye-tracker. The number of blinks per driving block (i.e., condition) was counted and normalised by the block duration (in s). We preprocessed the gaze signals by interpolating missing data (due to blinks and artefacts). Interpolation was done using a cubic spline method, with 50 ms padding around each gap and a maximum tolerated data loss of 500 ms (Kret & Sjak-Shie, 2019; Mathôt et al., 2018). The time series were then filtered using a median-based rolling window of 20 ms.

Calculation of Gaze Dispersion. We quantified gaze dispersion using the root mean square (*RMS*) distance of fixations from the gaze centroid within each driving block (i.e., condition). The gaze centroid was defined as the duration-weighted mean of the x- and y-coordinates across all fixations. Implementation details are provided in the Supplementary Material 8.5.

Calculation of Pupillary Measures. For the pupillary measures, pupil diameter was first converted to millimetres (see Lingelbach & Rieger, 2025, details can also be found in the Supplementary Material 8.5). Afterwards, we removed outliers below 1 mm and above 9 mm (outside the physiologically valid range; Kret & Sjak-Shie, 2019). Finally, the median pupil diameter during fixations was computed. Fixation events were obtained via the online event parser of the EyeLink eye-tracker.

The second pupillary measure IPA quantifies the fluctuations in pupil diameter as the frequency of abrupt discontinuities detected within the signal (in events per second; Duchowski, 2018). To compute the IPA, we applied a wavelet-based algorithm described in detail by Duchowski (2018), and in the Supplementary Material 8.5.

Inferential Statistics of Subjective, Behavioural, and Gaze-Related Correlates

For the subjective ratings, driving performance, and gaze-related measures, we used linear mixed-effects models (LMMs; Baayen et al., 2008), as implemented in the toolbox pymer4 (version 0.8.0; Jolly, 2018), to investigate main effects of the factors emotional speech (R.1; H.1) and visuo-spatial workload (R.1; H.2), as well as their interaction (R.2; H.3). Outliers exceeding the 95^{th} percentile were excluded and values rescaled using a z-score standardisation. Furthermore, the participant variable was incorporated as a random intercept in the models to account for non-systematic variations among individuals. Fixed effects were tested using *F*-statistics derived from Type-III sums of squares. If distributional assumptions were violated, the model was reparameterised and refitted using orthogonal polynomial contrasts before computing the ANOVA results. Degrees of freedom for fixed effects were estimated using the Satterthwaite approximation. Similar to the MEG analyses, post-hoc comparisons of significant effects were performed by using FDR-corrected Wilcoxon signed-rank tests, and visualising via bootstrapped means and their CIs (MCS with 5,000 iterations) of the contrasts (main effects; R.1 and H.1–2) and conditions (interaction effects; R.2 and H.3)

4.4 Results

4.4.1 Subjective Ratings

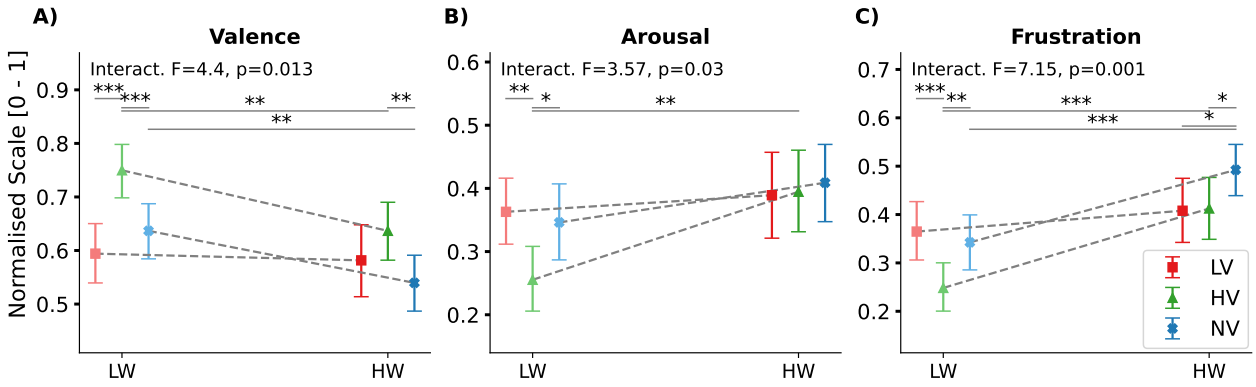
Linear mixed-effects models examined whether the perception of overall valence, arousal, effort, frustration, and distraction is co-modulated by emotional speech and visuo-spatial workload (R.2; H.3).

In line with this hypothesis, our results revealed that perceived valence, arousal, and frustration were co-modulated by the combination of emotional speech and visuo-spatial workload (interaction effect; valence: $F_{(2,\ 206.49)} = 4.40$, $p = .013$; excluded outlier observations: $11.63\ \%$; arousal: $F_{(2,\ 222.00)} = 3.57$, $p = .030$; excluded outlier observations: $5.11\ \%$; frustration: $F_{(2,\ 220.85)} = 7.150$, $p < .001$; excluded outlier observations: $5.49\ \%$). We did not observe a significant interaction effect on perceived effort and distraction.

Next, we investigated the nature of the significant interactions using post-hoc tests. As hypothesised, subjective ratings significantly differed between drives with positive and negative speech under low, but not high, workload (Figure 4.2). Increased workload diminished the perceived positivity of positive speech, which rendered responses to positive speech more similar to those elicited by negative speech. Neutral stimuli were perceived more negatively under high workload. Subjective experiences of negative speech were unaffected by the workload level.

Figure 4.2

Post-Hoc Comparisons of Interaction Effects Between Emotional Speech and Visuo-Spatial Workload in Ratings of Perceived (A) Valence, (B) Arousal, and (C) Frustration



Note. Coloured symbols and error bars represent the bootstrapped grand averages and their Bonferroni-corrected 2.5^{th} and 97.5^{th} confidence interval (CI) across participants. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload. Significance level from the Wilcoxon signed-rank test corrected with false discovery rate: *** for $p < .001$, ** for $p < .01$, * for $p < .05$.

Specifically, valence ratings were significantly lower in the negative compared to positive speech condition under low workload (LV – HV LW: $Z_{(47)} = -4.39$, $p < .001$; Figure 4.2A). The difference was

absent under high visuo-spatial workload (n.s.), due to a significant decrease in perceived valence for drives with positive speech (HV HW – LW: $Z_{(47)} = -3.55, p = .001$). For drives with neutral speech, workload significantly reduced perceived valence (NV HW – LW: $Z_{(47)} = -3.39, p = .002$).

Arousal and frustration ratings were significantly higher for drives with negative compared to positive speech under low workload (arousal LV – HV LW: $Z_{(47)} = -3.23, p = .006$; frustration LV – HV LW: $Z_{(47)} = -3.83, p < .001$; Figure 4.2B,C). Furthermore, low workload drives with positive speech were significantly lower rated in arousal than those with neutral speech (HV – NV LW: $Z_{(47)} = -2.58, p = .029$). From low to high workload, perceived arousal ($Z_{(47)} = -3.77, p = .002$) and frustration ($Z_{(47)} = -4.00$, $p < .001$) increased for drives with positive speech, eliminating any differences between the positive and negative speech conditions under high workload. Finally, perceived frustration during neutral speech increased with workload (NV HW – LW: $Z_{(47)} = -4.26, p < .001$), leading to significantly higher frustration for high workload drives with neutral compared to negative speech (LV – NV HW: $Z_{(47)} = -2.47, p = .021$).

In addition to these interaction effects, main effects of emotional speech and workload were observed for valence, arousal, frustration, distraction, and effort.

Due to the significant interactions on valence and frustration ratings, only the main effects of emotional speech in the contrast between HV and NV can be interpreted in these dependent variables (valence: $F_{(2,\ 207.76)} = 11.34, p < .001$; frustration: $F_{(2,\ 221,28)} = 9.72, p < .001$). The post-hoc tests indicate that drives with positive speech elicited higher valence (HV – NV: $M = 0.11, 95\ \%\ CI\ [0.07,\ 0.14]$; $Z_{(47)} = -4.23, p < .001$) and lower frustration ratings (HV – NV: $M = -0.09, 95\ \%\ CI\ [-0.12,\ -0.05]$; $Z_{(47)} = -4.24, p < .001$) than those with neutral speech, regardless of the current workload.

Contrary to our hypothesis that perceived effort and distraction should be co-modulated by the combination of emotional speech and workload, we did not observe significant interaction effects. The analyses, however, revealed that drives at high visuo-spatial workload and with negative and neutral speech were perceived as more effortful. Hence, the experienced workload was affected by the driving-related workload and emotional speech separately, but not by their combination. Moreover, emotional speech, particularly negative speech, was rated as more distracting.

This was reflected in significant main effects of emotional speech ($F_{(2,\ 206.97)} = 4.82, p = .009$; $11.63\ \%$ outliers excluded) and workload ($F_{(1,\ 207.24)} = 18.42, p < .001$; $11.63\ \%$ outliers excluded) on perceived effort, as well as a significant main effect of emotional speech on perceived distraction ($F_{(2,\ 195.89)} = 5.90, p = .003$; $11.57\ \%$ outliers excluded).

Post-hoc testing demonstrated that drives at high visuo-spatial workload were perceived as more effortful than at low workload (HW – LW: $M = 0.11, 95\ \%\ CI\ [0.07,\ 0.15]$; $Z_{(47)} = -4.48, p < .001$). In addition, listening to negative and neutral compared to positive speech was rated as more effortful (LV – HV: $M = 0.08$, $95\ \%\ CI\ [0.03,\ 0.13]$; $Z_{(47)} = -275, p = .009$; HV – NV: $M = -0.09, 95\ \%\ CI\ [-0.13,\ -0.04]$; $Z_{(47)} = -3.50, p = .001$; Figure 4.3A; H.1).

Negative speech was rated as more distracting than both neutral (LV – NV: $M = 0.12$, $95\ \%\ CI\ [0.04,\ 0.19]$; $Z_{(44)} = -2.88,\ p = .012$) and positive speech (LV – HV: $M = 0.07$, $95\ \%\ CI\ [0.01,\ 0.14]$; $Z_{(44)} = -2.22, p = .040$; Figure 4.3B; H.1). There was a marginal effect suggesting

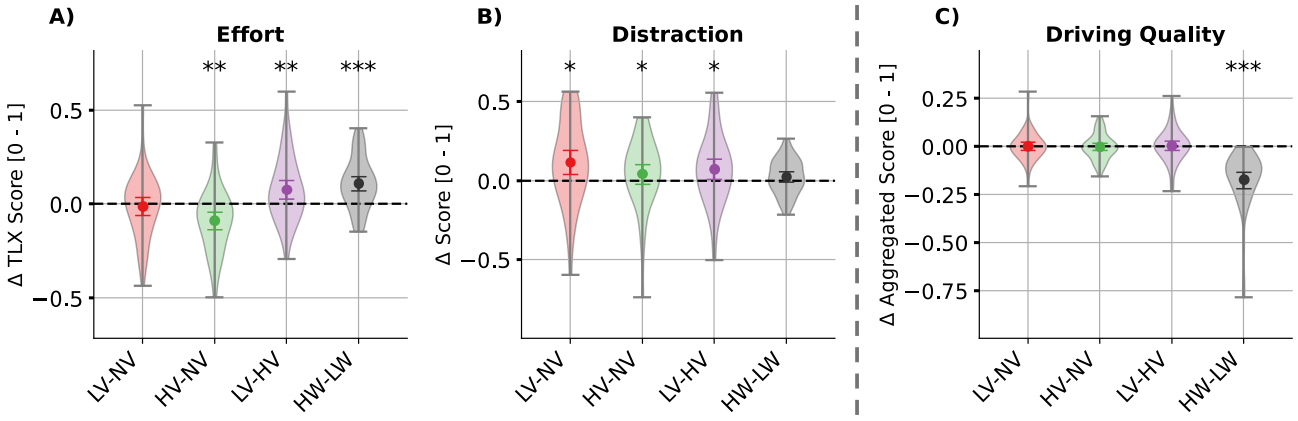
greater distraction from positive compared to neutral speech (HV – NV: $M = 0.04, 95\,\%\,CI\,[-0.02,\ 0.10]$; $Z_{(44)} = -1.96, p = .0499$).

An exploratory analysis using Spearman rank correlations (r_s) revealed that ratings of perceived effort and distraction, as well as speech arousal were highly correlated within participants (effort – distraction: $r_s = .459$, $p < .001$; effort – speech arousal: $r_s = .316$, $p < .001$; distraction – speech arousal: $r_s = .518$, $p < .001$; $df = 46$).

Detailed statistics on the subjective ratings are provided in the Supplementary Table 8.2, and Supplementary Figure 8.2, along with supplementary analyses on speech valence, speech arousal, and recall (Supplementary Table 8.3).

Figure 4.3

Post-Hoc Comparisons of Main Effects of Emotional Speech and Visuo-Spatial Workload in the Subjective Ratings of Perceived (A) Effort and (B) Distraction, as well as (C) Driving Quality.



Note. Coloured dots and error bars represent the bootstrapped grand averages and their Bonferroni-corrected 2.5th and 97.5th confidence interval (CI) across participants. The shaded area indicates the estimated probability density of the data. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload. Significance level from the Wilcoxon signed-rank test corrected with false discovery rate: *** for $p < .001$, ** for $p < .01$, * for $p < .05$.

4.4.2 Driving Performance

Results of the linear mixed-effects model for driving quality confirmed that it was more difficult to navigate through driving scenarios with high compared to low visuo-spatial workload (manipulation check). This was reflected in a significant main effect of visuo-spatial workload ($F_{(1,\ 215.53)} = 263.22, p < .001$; excluded outlier observations: $11.63\ \%$), and significantly lower driving quality under high compared to low workload (HW – LW: $M = -0.17, 95\,\%\,CI\,[-0.22,\ -0.13]$; $Z_{(47)} = -6.03, p < .001$; Figure 4.3C). We observed significant main effects of visuo-spatial workload in all subvariables of the driving quality score

(i.e., acceleration *RMSSD*, steering wheel angle *RMSSD*, mean brake actuation, mean deviation from lane centre, and number of traffic rule violations), supporting our decision to aggregate them into one score (see Supplementary Table 8.4 and its note for details). There were no differences between the emotional speech conditions, nor was there a significant interaction effect.

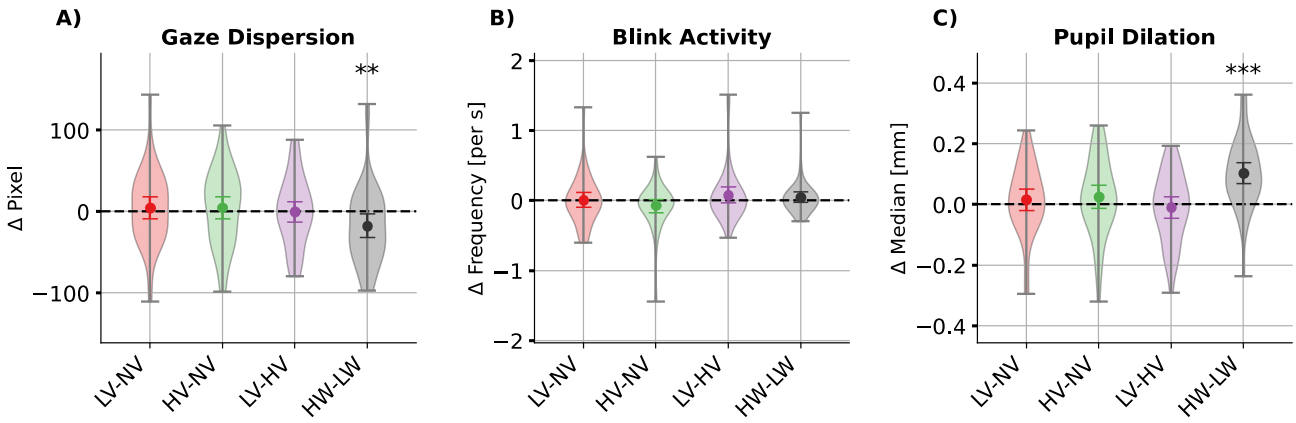
4.4.3 Gaze Behaviour

Next, we investigated how the different processing modes of emotional speech listening and visuo-spatial workload manifest in gaze behaviour (R.1). The linear mixed-effects models assessed whether the number of discontinuities in pupil size changes (IPA) increases during emotional speech (H.1), and whether driving complexity increases pupil dilation, and decreases blink frequency and gaze dispersion (H.2).

In line with the H.2, we observed a significant main effect of visuo-spatial workload on pupil dilation ($F_{(1,\ 182.88)} = 37.53$, $p < .001$; excluded outlier observations: 11.50 %), with larger pupil size during high workload (HW – LW: $M = 0.14$, $95\ \%\ CI\ [0.07,\ 0.10]$; $Z_{(42)} = -4.67$, $p < .001$; Figure 4.4C). In addition, gaze dispersion reduced ($F_{(1,\ 181.51)} = 9.95$, $p = .002$; excluded outlier observations: 11.50 %), indicating visual tunnelling (HW – LW: $M = -2.74$, $95\ \%\ CI\ [-31.33,\ -17.57]$; $Z_{(42)} = -2.82$, $p = .004$; Figure 4.4A).

Figure 4.4

Post-Hoc Comparisons of Main Effects of Emotional Speech and Visuo-Spatial Workload in the Gaze Behaviour Measures (A) Gaze Dispersion, (B) Blink Activity, and (C) Pupil Dilation



Note. Coloured dots and error bars represent the bootstrapped grand averages and their Bonferroni-corrected 2.5th and 97.5th confidence interval (CI) across participants. The shaded area indicates the estimated probability density of the data. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload. Significance level from the Wilcoxon signed-rank test corrected with false discovery rate: *** for $p < .001$, ** for $p < .01$, * for $p < .05$.

Contrary to our hypothesis, blink frequency was not significantly affected by increasing visuo-spatial workload in the driving task. Conversely, emotional speech did affect blink frequency ($F_{(2,\ 180.12)} = 4.53$,

$p = .012$; excluded outlier observations: $11.50\ \%$) and pupil dilation ($F_{(2,\ 179.35)} = 3.95, p = .021$; excluded outlier observations: $11.50\ \%$). However, the post-hoc tests did not reveal significant differences (Figure 4.4B). Moreover, emotional speech had no significant effect on the IPA (contrary to H.1). Descriptive statistics and a detailed statistical summary can be found in Supplementary Figure 8.3A – D and Supplementary Table 8.4, respectively.

4.4.4 Modulated Oscillatory Power With Dynamic Imaging of Coherent Sources (R.1)

We applied permutation-based spatial clustering (Maris & Oostenveld, 2007) to oscillatory source space power localised with DICS (J. Gross et al., 2001). We investigated oscillatory modulations in the alpha, beta, and gamma bands to examine distinct information processing modes during emotional speech and visuo-spatial cognition (R.1; main effects). Since we were interested in the main effects, we averaged the source band power across epochs of the respective other factor.

The permutation-based spatial clustering analyses revealed two distinct information processing modes associated with emotional speech and visuo-spatial workload, supporting H.1 and H.2. In line with a top-down guided processing during emotional speech (H.1), oscillatory beta-band power increased in parietal regions previously linked to predictive listening (Ou & Law, 2019) during drives with emotional relative to neutral speech. Concurrently, alpha-band power increased and gamma-band power decreased in parietal areas associated with driving-related spatial processing and attention (Fort et al., 2010; Sonnleitner et al., 2012). In line with a bottom-up processing during increased visuo-spatial workload (H.2), gamma-band power increased in the motor cortex and prefrontal regions involved in inhibitory control and decision-making (Fries, 2005; Kang et al., 2014). Concurrently, beta-band power was reduced in fronto- and temporo-parietal regions.

Main Effects of Emotional Speech

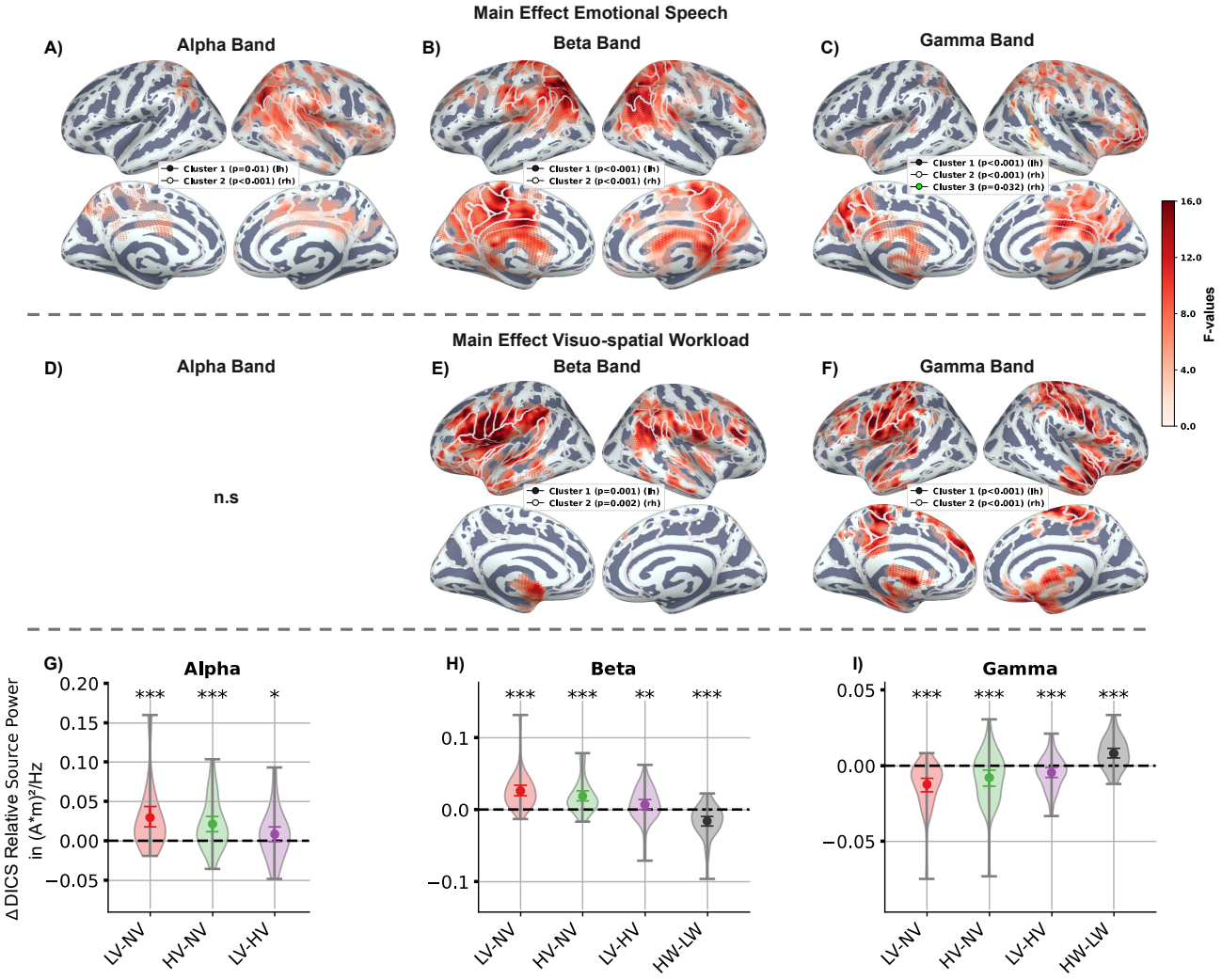
We found clusters yielding a significant main effect of emotional speech on modulations in alpha (*F*-statistic; left hemisphere (lh): $p = .01$; right hemisphere (rh): $p < .001$; Figure 4.5A), beta (lh and rh: $p < .001$; Figure 4.5B), and gamma-band power (lh and rh: $p < .001$; additional cluster 3 in rh: $p = .032$; Figure 4.5C).

The alpha-band cluster was primarily located in parietal regions implicated in spatial cognition and attention (Fort et al., 2010; Sonnleitner et al., 2012). The clusters comprised the superior parietal lobe, precuneus, right angular gyrus, and posterior cingulate cortex (Figure 4.5A). Beta-band clusters encompassed bilateral superior and inferior parietal lobes, the precuneus, intraparietal sulcus, angular gyrus, and cingulate cortex (Figure 4.5B). Gamma-band clusters included the bilateral precuneus and right OFC, with an additional cluster located near the posterior STS and right temporo-parietal junction (TPJ; Figure 4.5C).

To examine which conditions differed in their modulations in the clusters, oscillatory power was condition-wise averaged across vertices within significant clusters, and FDR-corrected Wilcoxon signed-rank tests were computed.

Figure 4.5

Results of the Permutation-Based Spatial Clustering of Source Power Localised with DICS



Note. The spatial distribution of significant *F*-values from the *F*-test clusters for the main effect of emotional speech is shown in Panel (A) for alpha-band power, (B) for beta-band power, and (C) for gamma-band power. The spatial distribution of significant *F*-values for the main effect of visuo-spatial workload is displayed in Panel (E) for beta-band power and (F) for gamma-band power. Results and contrasts of the post-hoc Wilcoxon signed-rank tests are shown in Panels (G) to (I). Band power was averaged across vertices within significant *F*-test clusters for each frequency band and effect. The shaded area indicates the estimated probability density of the data. Significance levels from the Wilcoxon signed-rank test, corrected using false discovery rate: *** for $p < .001$, ** for $p < .01$, * for $p < .05$. rh: right hemisphere, lh: left hemisphere. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

The post-hoc tests contrasting the speech conditions revealed increased alpha-band power during emotional compared to neutral speech (LV – NV: $M = 0.03$, 95 % CI $[0.02,\ 0.04]$; $Z_{(47)} = -6.53$, $p < .001$; HV – NV: $M = 0.02$, 95 % CI $[0.01,\ 0.03]$; $Z_{(47)} = -5.17$, $p < .001$), and a marginally

significant increase for negative compared to positive speech (LV – HV: $M = 0.01, 95\,\% CI\,[0.00,\ 0.02]$; $Z_{(47)} = -2.15, p = .031$; Figure 4.5G).

Similarly, beta-band power increased with emotional speech (LV – NV: $M = 0.03, 95\,\% CI\,[0.02,\ 0.03]$; $Z_{(47)} = -7.62, p < .001$; HV – NV: $M = 0.02, 95\,\% CI\,[0.01,\ 0.03]$; $Z_{(47)} = -6.69, p < .001$; LV – HV: $M = 0.01, 95\,\% CI\,[0.00,\ 0.01]$; $Z_{(47)} = -3.02, p = .003$; Figure 4.5H).

Conversely, gamma-band power decreased in response to emotional speech (LV – NV: $M = -0.01$, $95\,\% CI\,[-0.02,\ -0.01]$; $Z_{(47)} = -7.22, p < .001$; HV – NV: $M = -0.01, 95\,\% CI\,[-0.01,\ 0.00]$; $Z_{(47)} = -3.86, p < .001$), with a stronger decrease observed for negative compared to positive speech (LV – HV: $M = -0.01, 95\,\% CI\,[-0.01,\ 0.00]$; $Z_{(47)} = -3.47, p < .001$; Figure 4.5I).

Main Effects of Visuo-Spatial Workload

In the main effect of visuo-spatial workload, significant clusters were found in the beta (lh: $p = .001$; rh: $p = .002$; Figure 4.5E) and gamma band (lh and rh: $p < .001$; Figure 4.5F).

The beta-band cluster encompassed frontal regions such as the pars opercularis of the IFG, inferior dlPFC, frontal operculum, and anterior insula, as well as areas above the Sylvian fissure, including the inferior precentral gyrus and TPJ (Figure 4.5E). Gamma-band clusters covered the motor and parietal cortices, including the posterior cingulate cortex and precuneus, as well as the left frontal eye field, and mPFC and OFC (Figure 4.5F).

Post-hoc tests showed that gamma-band power increased in high compared to low visuo-spatial load drives (HW – LW: $M = 0.01, 95\,\% CI\,[0.00,\ 0.01]$; $Z_{(47)} = -4.11, p < .001$; Figure 4.5I), while beta-band power decreased (HW – LW: $M = -0.02, 95\,\% CI\,[-0.02,\ -0.01]$; $Z_{(47)} = -4.14, p < .001$; Figure 4.5H).

Supplementary Figure 8.7 presents the mean cluster band power per condition for the significant main effects. There was no significant interaction effect. Hence, all main effects can be interpreted.

Relationship Between Spatial Oscillatory Modulations

To investigate the relationship between significant markers of information processing, correlations were computed between averaged cluster band power per condition within participants. The correlation results suggest that efficient information processing during emotional speech and visuo-spatial workload occurs in tandem, with alpha- and beta-band modulations linked to oppositely directed gamma-band modulations (Engel et al., 2001).

During emotional speech, alpha- and beta-cluster band power were positively correlated ($r_s = .42$, $p < .001$; *df* $= 39$). Moreover, we observed negative correlations between gamma-cluster band power and alpha-cluster band power ($r_s = -.36, p < .001$; *df* $= 39$), as well as gamma- and beta-cluster band power ($r_s = -.28, p < .005$; *df* $= 39$).

In the workload conditions, gamma-cluster band power was significantly negatively correlated with beta-cluster band power ($r_s = -.51, p < .002$; *df* $= 39$). No other correlations reached significance for the main effect of visuo-spatial workload.

4.4.5 Multivariate Pattern Analysis of the Interaction Between Valence and Visuo-Spatial Workload

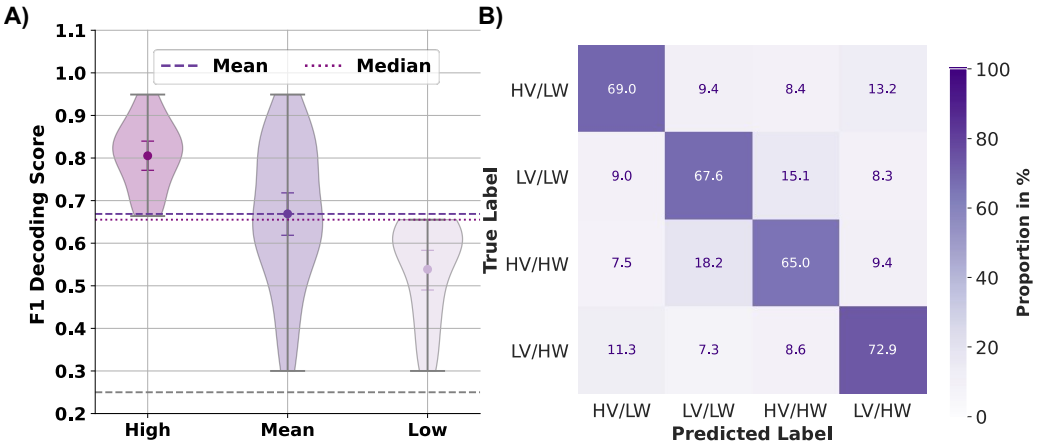
To address our second research question (R.2), we applied MVPA using CSP combined with LDA to discriminate drives with combinations of low and high visuo-spatial workload and negative and positive emotional speech in a four-class decoding task.

We assumed that the valence of speech and visuo-spatial workload co-modulate emotion appraisal and regulation. Our results in the subjective ratings supported this hypothesis, with emotional appraisal occurring only for positive speech under low workload (H.3). Under high workload, we expected that emotional appraisal would be inhibited and regulated. Motivated by findings from Kang et al. (2014) and Popov et al. (2012), which implicate prefrontal gamma as a mechanism of regulatory control, we investigated whether modulations in frontal gamma-band activity are linked to the regulation of emotional interference and can distinguish the four conditions.

As meta-analytic evidence indicates that individual differences strongly influence emotion regulation (Morawetz & Basten, 2024), we further analysed how low and high decoding performance in individuals was associated with personality and demographic characteristics.

Figure 4.6

Average Performance and Confusion Matrix of the Decoding



Note. (A) Mean *F*1 score across all subjects (mean; middle) and within the above median (high; left) and below median (low; right) decoding groups. Purple dashed line represents the mean, and dotted line the median. Dashed grey line represents the chance level at 0.25. Dots and error bars represent the bootstrapped grand averages, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI). (B) Confusion matrix displaying the proportion of models' predicted class label relative to the true class labels, averaged across folds and participants. LV: low valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Our results revealed clear discrimination among the four conditions in the test set, with performance well above the theoretical chance level (*F*1 score: 0.25), as indicated by an average *F*1 score of $M = 0.669$

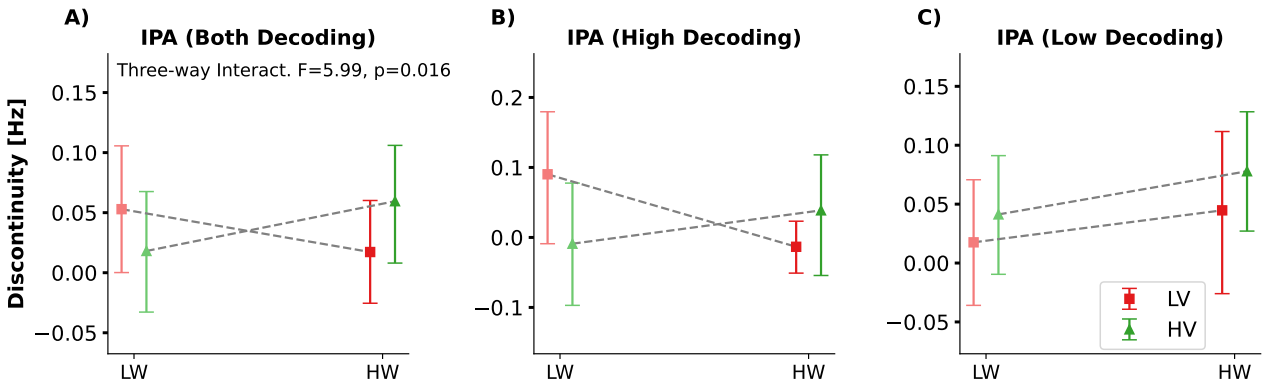
($95\,\%\,CI$: $[0.619,\ 0.718]$), and a median of 0.655 (range: 0.30 to 0.95). To investigate the relationship between decoding performance and interindividual differences, we divided the subjects into above-median (> 0.655) and below-median groups (≤ 0.655) based on decoding performance. The high-performing group had a mean *F*1 score of $M = 0.805$ ($95\,\%\,CI$: $[0.771,\ 0.840]$; $n = 20$), while the low-performing group had a mean of $M = 0.539$ ($95\,\%\,CI$: $[0.490,\ 0.583]$; $n = 21$). Figure 4.6A shows the mean *F*1 score and its CI across all participants (middle), as well as across those in the above-median (left), and below-median performance group (right; see Supplementary Figure 8.10 for the individual test decoding performances).

The averaged confusion matrix from the classifications revealed a cross-over interaction, with the most frequent misclassification occurring along the diagonal (Figure 4.6B). This observation already hints at an underlying cross-over interaction between valence and visuo-spatial workload.

The median-split analysis examined how the decoding performance groups differed in their psychophysiological traits, driving experience, and demographic characteristics. The analysis revealed only a significant difference in gender proportions across decoding groups (Fisher's exact test: Odds Ratio = 0.21; $p < .029$), with fewer female participants in the high decoding group (high decoding group: $n_{female} = 6$ and $n_{male} = 14$; low decoding group: $n_{female} = 14$ and $n_{male} = 7$).

Figure 4.7

Three-way Interaction Between Decoding Group, Valence of Speech, and Workload



Note. Coloured dots and error bars represent the bootstrapped grand averages and their Bonferroni-corrected $95\,\%$ confidence intervals per condition for the (A) full sample, (B) high decoding group (above median), and (C) low decoding group (below median). LV: low valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

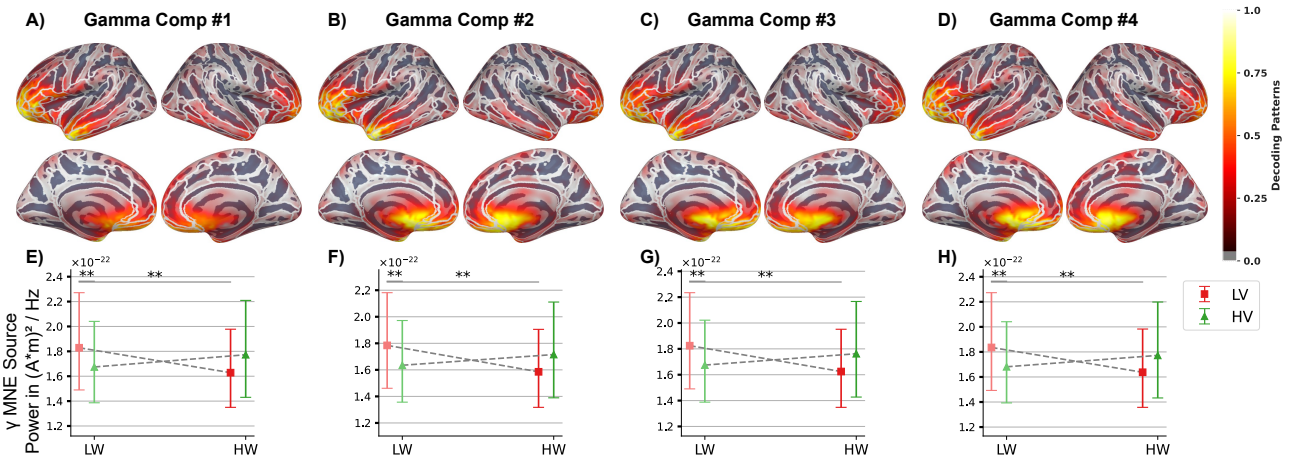
In exploratory analyses, we examined whether decoding group membership (i.e., the ability to decode individual neural signatures correctly) influenced the co-modulation of valence and workload on subjective measures (valence, arousal, frustration, effort, and distraction) and the IPA, as hypothesised in H.3. This was tested by including decoding group membership as a three-way interaction term in the LMMs. While no effects

of decoding group were found on subjective measures, a significant three-way interaction was observed for the IPA ($F_{(1,\ 92.15)} = 5.99, p = .016$; excluded outlier observations: $12.50\ \%$; Figure 4.7). In the high decoding group, IPA showed a valence-by-workload interaction: During low visuo-spatial workload, IPA increased for negative speech and decreased for positive speech. Conversely, under high workload, IPA increased for positive speech and decreased for negative speech. In contrast, the low decoding group showed no such interaction. In this group, IPA slightly increased with workload regardless of speech valence (LMM $95\ \%\ CI$ of the interaction effect: $[0.30,\ 2.95]$; bootstrapping with 5,000 iterations).

Localisation of Informative Neural Dynamics From CSP Components. To localise informative neural dynamics in the gamma band, the four CSP components were transformed into source space (Figure 4.8A – D). This analysis identified the IFG, OFC, anterior temporal lobe, and mPFC as key regions underlying the interaction between emotional speech processing and attentional control during visuo-spatial workload. The decoding patterns were more pronounced in the left hemisphere, and the components exhibited highly similar spatial distributions in source space (Figure 4.8A – D).

Figure 4.8

Decoding Patterns of the Four Common Spatial Pattern Components



Note. Components are used in the linear discriminant analysis (LDA). Panels (A) to (D) visualise the normalised values (range: 0 to 1) in the lateral (upper row) and medial view (lower row). Panels (E) to (H) display the source gamma-band power per condition for vertices exceeding the 90th percentile of the decoding pattern values. Significance levels from the Wilcoxon signed-rank test, corrected using false discovery rate: *** for $p < .001$, ** for $p < .01$, * for $p < .05$. LV: low valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

To examine the relationship between the decoding patterns and experimental conditions, we averaged source-level gamma-band power per condition across vertices exceeding the 90th percentile of the component

values (Figure 4.8E – H). Wilcoxon sign-rank tests revealed cross-over interactions, indicating a co-modulation between speech valence and visuo-spatial workload: During low visuo-spatial workload, source gamma-band power was significantly higher for drives with concurrent negative compared to positive speech (LV – HV LW – Component 1: $Z_{(40)} = -2.94, p = .005$; Component 2: $Z_{(40)} = -3.01, p = .004$; Component 3: $Z_{(40)} = -2.97, p = .005$; Component 4: $Z_{(40)} = -3.03, p = .004$). However, during high visuo-spatial workload, source gamma-band power significantly decreased for negative compared to positive speech (LV – HV HW – Component 1: $Z_{(40)} = -2.51, p = .015$; Component 2: $Z_{(40)} = -2.40, p = .021$; Component 3: $Z_{(40)} = -2.43, p = .019$; Component 4: $Z_{(40)} = -2.46, p = .018$).

While decoding increases sensitivity to effects affected by inter-individual neural variability, mass-univariate, permutation-based clustering is a commonly used complementary approach that identifies correlates in a data-driven manner at the group level (Maris & Oostenveld, 2007). To examine whether both approaches converge in their evidence, we performed a supplementary analysis with permutation-based clustering. This analysis revealed significant clusters of a cross-interaction in the gamma band (Supplementary Figure 8.12). The identified gamma-band clusters comprised the OFC, ACC, and mPFC (lh: $p = .03$; rh: $p = .012$; Supplementary Figure 8.12A). We performed post-hoc tests on gamma-band activity averaged across vertices within the significant cluster to compare the four conditions.

In line with our previous decoding results, the post-hoc tests revealed a significant cross-over interaction: During low workload, gamma power was significantly higher for negative speech than for positive speech (LV – HV LW: $Z_{(40)} = -3.18, p = .001$). Across workload levels, gamma power decreased for negative speech (LV HW – LW: $Z_{(40)} = -3.82, p < .001$) and increased for positive speech (HV HW – LW: $Z_{(40)} = -2.59$, $p = .009$). Consequently, during high workload, gamma power was significantly lower for negative compared to positive speech (LV – HV HW: $Z_{(40)} = -3.53, p < .001$; Supplementary Figure 8.12B).

Further supplementary analyses showed similar cross-interactions for the aperiodic broadband within the decoding patterns, and when not subtracting the aperiodic component from the gamma-band source power (Supplementary Figure 8.11 and 8.14).

4.5 Discussion

The MEG–eye-tracking study addressed two research questions: (1) What are the distinct neural signatures of emotional speech processing and visuo-spatial cognition during the dual task of driving and listening (R.1)?, and (2) do the valence of emotional speech and level of visuo-spatial workload jointly modulate information processing (R.2)?

To investigate these questions in a near-realistic scenario, we employed MEG, eye-tracking, a driving simulator, and naturalistic emotional speech stimuli from the validated GAUDIE database (Lingelbach et al., 2024). Using a dual-task paradigm involving driving under varying visuo-spatial workload and concurrent emotional speech listening, we hypothesised three distinct processing mechanisms. Emotional speech processing was expected to decrease gamma-band power in parietal regions and increase beta-band power associated with predictive listening (R.1). In contrast, increased visuo-spatial workload was hypothesised to

result in decreased beta-band and increased gamma-band power, along with increased pupil dilation (R.1). A third, regulatory processing mode was postulated to reflect emotion regulation through frontal gamma-band activity. Specifically, we hypothesised that positive emotional speech would be attended to and appraised under low, but not high, visuo-spatial workload. Under high workload, gamma-band-based regulatory activity was expected to suppress emotional interference to support goal-directed processing associated with the driving task (R.2).

Our findings revealed distinct neural, gaze-related, and behavioural signatures distinguishing antagonistic information processing modes for emotional speech processing and visuo-spatial cognition (R.1). In addition, subjective ratings indicated that speech valence and visuo-spatial workload co-modulated both inhibitory control and emotional appraisal of positive speech (R.2). Specifically, emotional appraisal of positive speech occurred only under low, but not high, visuo-spatial workload. At the neural level, multivariate pattern analysis revealed a co-modulatory effect of speech valence and visuo-spatial workload on fronto-temporal gamma-band activity (R.2). Gamma oscillations in the OFC, temporal pole, and mPFC indicated reduced inhibitory control for positive compared to negative speech under low workload, and increased inhibitory control for positive compared to negative speech under high workload. Negative speech triggered the regulatory mode even during low workload; however, the effectiveness of top-down control diminished as workload increased.

The study offers converging evidence for distinct processing modes underlying emotional speech processing, visuo-spatial cognition, and regulatory control mechanisms that inhibit emotional interference as a function of available cognitive resources in complex, naturalistic environments. The findings also revealed circumstances under which emotional speech valence and visuo-spatial workload jointly influence cognitive resource allocation and regulatory mechanisms.

4.5.1 Oscillatory Dynamics Guiding Information Processing in Multisensory Environments

Emotional Speech Processing Involves Predictive Listening and Social Cognition

Consistent with our hypothesis on the role of beta- and gamma-band oscillations during emotional speech listening, we found that increased parietal beta- and decreased gamma-band power served as distinct markers of emotional compared to neutral speech processing. Additionally, we observed modulations in alpha-band power in temporo-parietal regions.

The spatial distribution of the beta-band cluster (Figure 4.5B) is in concordance with the previously described dorsal stream of deliberate emotion processing (Phillips et al., 2003; Tucker et al., 1995). We did not observe significant clusters in ventral stream areas associated with sensory processing and automatic emotion processing (Phillips et al., 2003; Schirmer & Kotz, 2006; Tucker et al., 1995).

One possible explanation is that the dual task, characterised by distributed cognitive resources, may have interfered with low-level acoustic analysis involved in automatic emotion recognition (ventral processing stream; Schirmer & Kotz, 2006), instead engaging dorsal regions associated with deliberate emotion processing and later evaluative stages (Phillips et al., 2003; Tucker et al., 1995). Previous linguistic research has associated increased parietal beta activity with top-down mechanisms of deliberate anticipation and prediction in speech

perception (Hovsepyan et al., 2023; Ou & Law, 2019). Although speech predictability was not directly manipulated in our study, the oscillatory patterns observed during emotional compared to neutral speech processing resemble those reported by Hovsepyan et al. (2023) and Ou and Law (2019), suggesting that socio-emotional speech may facilitate predictive listening. In contrast, purely informative, neutral speech, such as weather forecasts, appears to be less conducive to prediction. In their electroencephalography study, Pinheiro et al. (2017) reported enhanced predictive processing reflected in increased beta-band activity exclusively for positive but not negative vocalisations. In contrast, our findings reveal such enhancement for both positive and negative speech sequences. Importantly, Pinheiro et al. (2017) employed brief, context-free vocalisations, such as laughter and growls. Such stimuli may have complicated predictions and elicited weaker effects than coherent, naturalistic speech sequences (Hagoort, 2019; Hamilton & Huth, 2020).

Alternatively, predictive listening during socio-emotional speech processing could also be related to the theory of mind (ToM). Cognitive ToM refers to the ability to represent one's own and others' thoughts, intentions, and desires (Schurz et al., 2021), and has been associated with activation in the precuneus, cuneus, the middle and superior temporal gyri (Schlaffke et al., 2015; Schurz et al., 2021), and right TPJ (Saxe & Kanwisher, 2003; Schurz et al., 2021). This functional network mirrors the spatial distribution of the observed beta-band clusters in the main effect of emotional speech (Figure 4.5B).

In addition to the beta-band clusters, we observed clusters in the alpha and gamma bands for the main effect of emotional speech. These encompassed parietal regions, including the superior parietal lobe, precuneus, and cuneus, as well as the posterior cingulate cortex and posterior STS. They have previously been identified as the network relevant to driving, particularly for spatial attention and processing (Haghani et al., 2021; Palmiero et al., 2019). In agreement with previous studies studying speech processing during driving (Fort et al., 2010; Sasai et al., 2016; Sonnleitner et al., 2012), we propose that reduced gamma-band power, together with increased alpha-band power in this driving network, reflects diminished spatial cognition. This implies that cognitive resources are reallocated from driving-related processing towards emotional speech processing.

The subjective ratings and correlation analyses also support this interpretation. Emotional speech stimuli increased distractions. Likewise, higher speech arousal, also characterising the emotional speech stimuli, was correlated with greater effort and distraction during the driving task.

To summarise, listening to emotional speech likely triggered both enhanced anticipatory listening and socio-emotional cognition. At the same time, visuo-spatial attention and processing decreased. We propose that these are signatures of a top-down guided processing mode during emotional speech listening (resembling an internalised processing; Oosterwijk et al., 2015).

Spatial Cognition and Sensorimotor Processing Increase During Difficult Drives

Our results indicate that increased visuo-spatial workload elicits enhanced stimulus-driven bottom-up processing. Under high visuo-spatial workload and increased task demand, we observed decreased driving quality as well as increased subjective effort, pupil dilation, and visual tunnelling, reflected in reduced gaze dispersion (Nunes & Recarte, 2002, confirming H.2).

We proposed that modulations in pupil diameter index attentional effort during visuo-spatial attention and the tracking of multiple objects. Translational animal, stimulation, and neuroimaging studies have demonstrated that pupil dilation is linked to neural circuits involved in attention and visuo-spatial processing (Alnæs et al., 2014; Joshi et al., 2016; Vinck et al., 2015). These circuits include the nucleus locus coeruleus and associated noradrenergic system (Alnæs et al., 2014; Joshi et al., 2016; Vinck et al., 2015), the oculomotor system including the frontal eye fields and pathways to the superior colliculus (Alnæs et al., 2014; Joshi et al., 2016), cingulate cortex (Joshi et al., 2016), and superior parietal lobule (Alnæs et al., 2014).

At the neural level, we observed increased gamma-band power in parietal, motor, frontal, and fronto-temporal cortices during high compared to low visuo-spatial workload, while beta-band power decreased in regions dorsal to the Sylvian fissure (H.2). The spatial distribution of the gamma- and beta-band clusters partially overlaps with neural circuits linked to pupil dilation and aligns with findings from previous driving studies. These studies reported increased activation in motor and parietal regions during challenging driving scenarios, suggesting that these areas contribute to motor control, sensorimotor integration, and visuo-spatial processing (Navarro et al., 2018; Sasai et al., 2016; Spiers & Maguire, 2007). The antagonistic relationship between gamma- and beta-band power in the motor cortex and related areas, such as the frontal eye fields, has been well documented in electrocorticography (ECoG; e.g., Crone et al., 1998), stimulation (e.g., Joundi et al., 2012), and animal studies (e.g., Sendhilnathan et al., 2017).

Shi et al. (2023) offered a further interpretation of the role of gamma-band oscillations during challenging driving scenarios. The authors outlined in their review that OFC gamma-band activity is associated with the self-monitoring of decision-making in social contexts (Shi et al., 2023). Since the high visuo-spatial workload condition included unpredictable car agents, increased prefrontal gamma activity likely reflects driving-related decision-making (e.g., lane changes), as well as social evaluation and predictive processes in response to the other agents. Beta-band power, which was negatively correlated with gamma-band power, decreased in temporal and parietal regions, including the STG, right TPJ, and right supramarginal gyrus. This finding is in line with an MEG study of van Pelt et al. (2016). Using Granger causality, van Pelt et al. (2016) investigated the relationship between antagonistically operating beta- and gamma-band oscillatory activity in the context of predictive coding during the observation of animated bowling actions. In their study, oscillatory power in the TPJ and mPFC increased in the gamma band and decreased in the beta band during prediction errors. The authors proposed that predictive coding involves bottom-up sensory encoding, driven by gamma-band modulations, and top-down prediction evaluation, driven by beta-band modulations (van Pelt et al., 2016).

In summary, we observed oscillatory signatures indicative of stimulus-driven, bottom-up processing and the reallocation of cognitive resources to the driving task under high visuo-spatial workload. The gamma–beta oscillatory pattern suggests compensatory processes involving predictive cognition and motor responses. These compensatory processes are likely recruited to meet increased task demands and associated cognitive strain, as also indicated by greater pupil dilation.

Emotional Regulation Through Fronto-Temporal Gamma Oscillatory Activity is Co-Modulated by Valence and Workload

Our second research question and last hypothesis (H.3) target the co-modulation of valence and workload (i.e., interaction effects). We hypothesised that appraisal and processing of positive emotion and, thus, valence-specific signatures, would occur only when sufficient cognitive resources are available under low visuo-spatial workload (Baddeley, 1992; Cowan, 2017; Lingelbach & Rieger, 2025; Wickens, 2014). Under high visuo-spatial workload, prefrontal gamma oscillatory activity was expected to regulate emotional interference (Kang et al., 2014; Popov et al., 2012; Quirk & Beer, 2006).

Our findings largely supported the hypothesis and offered even further insights. A neural signature in the form of prefrontal and fronto-temporal gamma-band activity was co-modulated by the valence of speech and workload (cross-over interactions). In line with previous studies (Kang et al., 2014; Popov et al., 2012), we propose that this fronto-temporal signature is related to the regulatory processing of emotional interference. Under low workload, fronto-temporal gamma-band activity was greater for negative than positive speech (in line with H.3). Under high visuo-spatial workload, fronto-temporal oscillatory gamma activity increased for positive speech, and decreased for negative speech (reversed pattern).

The gamma-band signature included the OFC, temporal pole, mPFC, and ACC. These regions were previously linked to top-down guided cognitive control (Koechlin et al., 2003) and emotion regulation (see Quirk & Beer, 2006, for review; see Berboth & Morawetz, 2021; and Morawetz & Basten, 2024, for meta-analyses). The spatial distribution of the gamma-band signature partially overlaps with that observed under increased visuo-spatial workload, suggesting that these regions also contribute more generally to the coordination of goal-directed processes (Koechlin et al., 2003).

Emotion regulation describes the capacity to control and modify emotional processing through implicit or explicit strategies, such as distraction, reappraisal, and suppression (J. J. Gross, 1998). Using intracranial electroencephalography (iEEG) and an emotional picture viewing task, Sonkusare et al. (2023) proposed gamma activity in the mPFC and OFC to regulate negative emotions (see also Golkar et al., 2012; Popov et al., 2012; Quirk & Beer, 2006). Functional MRI studies indicate that emotion regulation by the mPFC, vlPFC, OFC, and temporal pole is likely achieved through inhibition of limbic emotion-related processing in the insula and amygdala (Berboth & Morawetz, 2021; Goldin et al., 2008; Golkar et al., 2012; Ochsner et al., 2012; Sonkusare et al., 2020). Lesion studies also reported that damage to the OFC impairs the ability to modulate or inhibit neural responses to aversive stimuli (Hooker & Knight, 2006; Rule et al., 2002). Moreover, the OFC and mPFC are involved in decision-making, self-reflection, and cognition in social contexts (Amodio & Frith, 2006; Bechara et al., 2000; see van Overwalle, 2009, for meta-analysis).

In light of these previous findings (e.g., Kang et al., 2014; Morawetz & Basten, 2024; Popov et al., 2012; van Overwalle, 2009), our results suggest that increased fronto-temporal gamma-band activity contributes to the inhibition of emotional responses to negative speech under low workload (see Lingelbach et al., 2023, for similar findings when using continuous naturalistic speech; but cf. Mothes-Lasch et al., 2012, using an event-related study with negative words). However, positive speech was emotionally perceived and appraised under low workload. The subjective ratings in our study support this interpretation, revealing higher valence, as

well as lower arousal and frustration for positive relative to negative speech, but only under low visuo-spatial workload.

Under high workload, gamma-band activity increased in response to positive speech. In line with H.3, we interpret this as active emotion regulation of positive speech due to now limited available processing resources. However, contrary to our expectations, fronto-temporal gamma activity decreased for negative speech at high workload. This finding has two potential implications: First, regulating negative speech appears to demand greater cognitive resources (Cromheeke & Mueller, 2014; Lingelbach & Rieger, 2025; Schweizer et al., 2019). Second, at high workload, these resources seem to become depleted, thereby compromising inhibitory control (Cowan, 2017; Wickens, 2014). Despite this, subjective ratings and driving behaviour in response to negative speech remained unaffected by workload.

Interestingly, more female participants revealed below-median decoding performance when classifying the four valence-workload conditions. This observation points to potential gender-specific differences in the integration of emotional and cognitive processes. Derntl et al. (2010) investigated gender differences in neural signatures during socio-emotional cognition and found that females and males relied on divergent processing strategies and neural networks. While females showed increased activation in subcortical limbic regions associated with emotional reactivity (signals that are not captured by MEG; Hämäläinen et al., 1993), males displayed increased activation in cortical regions associated with cognitive aspects of social cognition (Derntl et al., 2010).

Gaze Signatures Sensitive to the Emotion-Cognition Interaction. In line with Vogels et al. (2018), we observed clear distinctions between modulations in pupil diameter and discontinuities in pupil diameter changes. This suggests that these gaze-related measures may index different cognitive processes. While pupil dilation increased under high visuo-spatial workload (confirming H.2), discontinuities in pupil diameter changes remained unaffected. Contrary to our hypothesis that discontinuities in changes would be jointly modulated by speech valence and workload level, we found no significant co-modulation. However, exploratory analysis indicated that discontinuities in pupil diameter changes may be partially linked to the fronto-temporal gamma-band signature involved in regulating emotional interference. This preliminary observation warrants further investigation. To summarise, our findings support the notion that the pupillary measures index distinct cognitive processes. While pupil dilation appears to reflect attentional effort and the required cognitive resources to perform a task (van der Wel & van Steenbergen, 2018), the frequency of pupillary discontinuities may relate to attentional control involved in allocating or redistributing cognitive resources (Demberg & Sayeed, 2016; Vogels et al., 2018).

Complementary Approach of Multivariate Pattern and Mass-Univariate Analysis. Mass-univariate statistics such as permutation-based clustering examine activation from clusters separately (Haufe et al., 2014), and assume that neural activation patterns are similarly localised across subjects (Holdgraf et al., 2017; Marsicano et al., 2024). Consequently, they may be substantially limited when investigating individual-specific processing strategies (Marsicano et al., 2024) or naturalistic stimulus material (Sonkusare et al., 2019).

MVPA enhances sensitivity by integrating measurement sources within a multidimensional framework, while accounting for inter-individual neural variability (Haufe et al., 2014; Holdgraf et al., 2017; Marsicano et al., 2024). Our complementary approach, incorporating MVPA, provided novel insights into the role of fronto-temporal gamma-band oscillatory activity in the integration and interaction of emotional speech processing and visuo-spatial cognition.

In summary, our findings suggest that participants attended and appraised positive speech but only during low visuo-spatial workload. They aimed to inhibit negative speech regardless of the current workload level. However, the effectiveness of the top-down guided emotional regulation of negative speech declined as workload increased. This observation has important practical implications: As workload increases further, listening to negative speech may have a greater detrimental impact on driving performance than positive speech.

4.5.2 Studying the Brain in its Natural Environment

We employed a multimodal approach, combining eye-tracking and MEG, alongside a near-realistic experimental paradigm, to investigate effects of and interactions between cognitive and emotional processing in a driving simulator study. The experimental dual-task paradigm incorporated both a driving simulation, which mirrored authentic driving experiences, and naturalistic speech. In contrast to the only other available German emotional speech database (which includes 10 short sentences spoken with varying emotional prosody by professional actors; Burkhardt et al., 2005), the GAUDIE database features longer conversational speech with coherent context. Furthermore, it was developed and validated to explicitly induce positive, neutral, and negative emotions. Consequently, its emphasis lies on *how* participants emotionally experience the stimulus (Lingelbach et al., 2024), rather than solely on *what* valence the stimulus possesses (Burkhardt et al., 2005).

Naturalistic, ecologically valid research is expected to enhance transferability to real-world contexts, as it involves a lower degree of abstraction than studies using non-natural, isolated stimuli (Hamilton & Huth, 2020). Previous comparisons of linguistic studies have revealed processing differences between natural and isolated speech in terms of lateralisation and specialisation. Furthermore, our results suggest that continuous negative speech is already inhibited during low workload (see also Lingelbach et al., 2023), whereas studies using single words presented in an event-related design did not report emotion inhibition for negative words during low load (Mothes-Lasch et al., 2012).

According to the criteria defined by Hamilton and Huth (2020), our study demonstrates a high level of ecological validity. A German online survey of 1,700 participants indicated that $67\ \%$ regularly listen to audio content while travelling, covering $82\ \%$ of their total travel time when driving (Landesanstalt für Medien NRW, 2021). This underscores both the representativeness of the chosen paradigm and the practical relevance of understanding brain function and cognition during the dual task of driving while listening to speech or engaging in emotionally loaded in-car conversations. In real-life driving situations, additional sources of emotional responses, such as dangerous or rule-breaking manoeuvres performed by other drivers, can further impact cognitive processes. Therefore, future research should examine how different types of emotionally charged

events affect the integration of emotional and cognitive processes, as well as the regulation of emotional responses during driving.

4.5.3 Limitations

Several shortcomings of this study should be noted. Although we chose a fully equipped driving simulator with authentic driving scenarios, the experience of driving in the MEG remains considerably different from real-world driving. To minimise muscle artefacts, participants' upper body and arm movements were restricted with pillows during steering, and braking was performed with the left foot. Future studies employing EEG should aim to replicate the cortical signatures observed here in real-world, out-of-laboratory driving settings. Such EEG studies should further investigate whether midfrontal theta-band oscillatory activity, related to executive functioning during driving (Wascher et al., 2018), is co-modulated by emotional speech and visuo-spatial workload. While midfrontal theta is reliably detected with EEG, these signals primarily originate from superficial radial dipole layers, to which MEG is less sensitive (Srinivasan et al., 2006).

We observed a co-modulation of speech valence and workload on fronto-temporal gamma oscillations, reflected in a significant cross-over interaction in the clustering analysis when localising oscillatory source power using MNE (R.2), but not when using DICS (R.1). One possible explanation for the discrepancy between the two localisation methods is that DICS may have partially filtered out oscillatory changes related to emotion regulation (J. Gross et al., 2001; van Vliet et al., 2018), if participants had already engaged inhibitory control during baseline periods to detach from preceding emotional experiences.

Although naturalistic speech offers a more holistic emotional experience, disentangling the individual contributions of semantic content, paralinguistic features, and their interaction remains challenging in our study (but cf. Alday, 2019, for study approaches with natural stimuli). Hence, they likely contributed to the observed neurophysiological signatures during the dual task. Future research could examine oscillatory modulations and magnetic fields time-locked to salient events to further elucidate the interaction between emotional and cognitive processes in naturalistic experimental paradigms (i.e., salient acoustic events, Khalighinejad et al., 2017; but also events linked to eye blinks, Alyan et al., 2023; and object pursuits, Agtzidis et al., 2020).

4.6 Conclusion

In the study, we identified antagonistic neural processing modes for emotional speech and visuo-spatial cognition during the dual task of speech listening and driving, based on permutation-based clustering of oscillatory source power. Increased beta- and decreased gamma-band power in parietal regions indicated an internalised processing mode for emotional speech. This mode is likely associated with enhanced anticipatory listening, social cognition, and reduced spatial processing. In contrast, decreased beta-band and increased gamma-band source power, along with increased pupil dilation, signalled an externalised, bottom-up processing mode under high visuo-spatial workload. This likely reflects enhanced sensorimotor processing and spatial cognition, but also increased cognitive strain due to heightened task demands.

Multivariate pattern analysis and subjective ratings indicated emotional appraisal of positive speech

only under low, but not high, visuo-spatial workload. Under high visuo-spatial workload and reduced cognitive resources, gamma-band activity in the OFC, temporal pole, and mPFC appears to mediate the regulation of emotional speech. Participants attempted to inhibit negative speech regardless of workload level; however, the effectiveness of top-down emotional regulation of negative speech declined as workload increased.

To conclude, by combining neural, gaze, and behavioural correlates, we found converging evidence of distinct information processing modes for visuo-spatial cognition and auditory emotional speech processing in naturalistic multisensory environments. Lastly, our multivariate analyses revealed a gamma-based top-down regulatory signature in fronto-temporal regions. This mechanism flexibly up- and down-regulates emotional speech processing in response to potential audio-visual interference in complex, naturalistic environments.

Funding

The research was supported by the Fraunhofer Gesellschaft with the scholarship ”Fraunhofer TALENTA” (to K. Lingelbach), by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; RI 1511/3-1 awarded to J. Rieger), and by the Neuroimaging Unit of the Carl von Ossietzky University of Oldenburg, funded by DFG grants (3T MRI INST 184/152-1 FUGG and MEG INST 184/148-1 FUGG).

Declaration of Competing Interests

The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Data and Code Availability, Pre-Registration, Ethics Approval, and Consent of Participation and Publication

The study was pre-registered on the OSF (https://osf.io/bta49) and the protocol was approved by the ethics committee of the Carl von Ossietzky University, Oldenburg, Germany (Ref. EK/2018/070). Participants provided informed consent following the Declaration of Helsinki. They were informed that their participation was voluntary, that they could withdraw at any time during the experiment, and that they could request the deletion of their data up until the point of anonymisation. The informed consent also included a section on the publication of anonymised data and aggregated results. Data and code supporting the findings of this manuscript are available upon request via the OSF project (https://osf.io/um6vw/).

Acknowledgements

We thank Dr. Tina Schmitt for her support with the source reconstruction of the anatomical T1 scans, Dr. Andreas Spiegler for his assistance with the technical setup and data collection, and Jooris Mettler for his

support with the driving simulation. We would also like to thank Prof. Dr. Pascal Fries for his contributions to discussions on the interpretation of oscillatory activity and aperiodic components.

References Study III

Agtzidis, I., Meyhöfer, I., Dorr, M., & Lencer, R. (2020). Following Forrest Gump: Smooth pursuit related brain activation during free movie viewing. *NeuroImage*, *216*, 116491. https://doi.org/10.1016/j.neuroimage.2019.116491

Alday, P. M. (2019). M/EEG analysis of naturalistic stories: A review from speech to language processing. *Language, Cognition and Neuroscience*, *34*(4), 457–473. https://doi.org/10.1080/23273798.2018.1546882

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4), 1. https://doi.org/10.1167/14.4.1

Alyan, E., Arnau, S., Reiser, J. E., Getzmann, S., Karthaus, M., & Wascher, E. (2023). Blink-related EEG activity measures cognitive load during proactive and reactive driving. *Scientific Reports*, *13*(1), 19379. https://doi.org/10.1038/s41598-023-46738-0

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277. https://doi.org/10.1038/nrn1884

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., De Weerd, P., Kennedy, H., & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390–401. https://doi.org/10.1016/j.neuron.2014.12.018

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292. https://doi.org/10.1037/0033-2909.91.2.276

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, *10*(3), 295–307. https://doi.org/10.1093/cercor/10.3.295

Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *9*(9), 1395–1403. https://doi.org/10.1093/scan/nst124

Berboth, S., & Morawetz, C. (2021). Amygdala-prefrontal connectivity during emotion regulation: A meta-analysis of psychophysiological interactions. *Neuropsychologia*, *153*, 107767. https://doi.org/10.1016/j.neuropsychologia.2021.107767

Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., & Müller, K.-R. (2008). Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Processing Magazine*, *25*(1), 41–56. https://doi.org/10.1109/MSP.2008.4408441

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Briesemeister, B. B., Tamm, S., Heine, A., & Jacobs, A. M. (2013). Approach the good, withdraw from the bad – A review on frontal alpha asymmetry measures in applied psychological research. *Applied Psychology: An International Review*, *62*(4), 412–439. https://doi.org/10.1111/j.1464-0597.2012.00450.x

Briggs, G. F., Hole, G. J., & Land, M. F. (2011). Emotionally involving telephone conversations lead to driver error and visual tunnelling. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*(4), 313–323. https://doi.org/10.1016/j.trf.2011.02.004

Brockhoff, L., Schindler, S., Bruchmann, M., & Straube, T. (2022). Effects of perceptual and working memory load on brain responses to task-irrelevant stimuli: Review and implications for future research. *Neuroscience & Biobehavioral Reviews*, *135*, 104580. https://doi.org/10.1016/j.neubiorev.2022.104580

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., Weiss, B., et al. (2005). A database of German emotional speech. *Interspeech*, *5*, 1517–1520. https://doi.org/10.21437/Interspeech.2005-446

Chaumon, M., Bishop, D. V., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*, *250*, 47–63. https://doi.org/10.1016/j.jneumeth.2015.02.025

Chavarriaga, R., Ušćumlić, M., Zhang, H., Khaliliardali, Z., Aydarkhanov, R., Saeedi, S., Gheorghe, L., & Millán, J. d. R. (2018). Decoding neural correlates of cognitive states to enhance driving experience. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *2*(4), 288–297. https://doi.org/10.1109/TETCI.2018.2848289

Clayton, M. S., Yeung, N., & Kadosh, R. C. (2015). The roles of cortical oscillations in sustained attention. *Trends in Cognitive Sciences*, *19*(4), 188–195. https://doi.org/10.1016/j.tics.2015.02.004

Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, *18*(2), 192–205. https://doi.org/10.1097/00004728-199403000-00005

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cromheeke, S., & Mueller, S. C. (2014). Probing emotional influences on cognitive control: An ALE meta-analysis of cognition emotion interactions. *Brain Structure & Function*, *219*(3), 995–1008. https://doi.org/10.1007/s00429-013-0549-z

Crone, N. E., Miglioretti, D. L., Gordon, B., & Lesser, R. P. (1998). Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band. *Brain: A Journal of Neurology*, *121*(12), 2301–2315. https://doi.org/10.1093/brain/121.12.2301

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *The American Psychologist*, *60*(2), 170–180. https://doi.org/10.1037/0003-066X.60.2.170

Czerniak, J. N., Schierhorst, N. J., Brandl, C., Mertens, A., Schwalm, M., & Nitsch, V. (2021). A meta-analytic review of the reliability of the index of cognitive activity concerning task-evoked cognitive workload and light influences. *Acta Psychologica*, *220*, 103402. https://doi.org/10.1109/TETCI.2018.2848289

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Demberg, V., & Sayeed, A. (2016). The frequency of rapid pupil dilations as a measure of linguistic processing difficulty. *PLOS ONE*, *11*(1), 1–29. https://doi.org/10.1371/journal.pone.0146194

Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D. I., Schneider, F., & Habel, U. (2010). Multidimensional assessment of empathic abilities: Neural correlates and gender differences. *Psychoneuroendocrinology*, *35*(1), 67–82. https://doi.org/10.1016/j.psyneuen.2009.10.006

Dolcos, F., Iordan, A. D., & Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: A review of evidence from brain imaging investigations. *Journal of Cognitive Psychology*, *23*(6), 669–694. https://doi.org/10.1080/20445911.2011.594433

Dong, X., Cui, L., & Johnson, B. W. (2024). Neural mechanisms for secondary suppression of emotional distractors: Evidence from concurrent electroencephalography–magnetoencephalography data. *Emotion*, *24*(8), 1907–1922. https://doi.org/10.1037/emo0001388

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, *23*(12), 1655–1665. https://doi.org/10.1038/s41593-020-00744-x

Duchowski, A. T. (2018). The index of pupillary activity: Measuring cognitive load vis-à-vis task difficulty with pupil oscillation. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, *282*, 1–13. https://doi.org/10.1145/3173574.3173856

Engel, A. K., & Fries, P. (2010). Beta-band oscillations – Signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165. https://doi.org/10.1016/j.conb.2010.02.015

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716. https://doi.org/10.1038/35094565

Engemann, D. A., & Gramfort, A. (2015). Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals. *NeuroImage*, *108*, 328–342. https://doi.org/10.1016/j.neuroimage.2014.12.040

European Commission & Directorate General for Mobility and Transport. (2020). *Next steps towards 'Vision Zero' – EU road safety policy framework 2021-2030*. Publications Office of the European Union. https://doi.org/10.2832/391271

Faure, V., Lobjois, R., & Benguigui, N. (2016). The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, *40*, 78–90. https://doi.org/10.1016/j.trf.2016.04.007

Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *8*(4), 272–284. https://doi.org/10.1002/(sici)1097-0193(1999)8:4⟨272::aid-hbm10⟩3.0.co;2-4

Fort, A., Martin, R., Jacquet-Andrieu, A., Combe-Pangaud, C., Foliot, G., Daligault, S., & Delpuech, C. (2010). Attentional demand and processing of relevant visual information during simulated driving: A MEG study. *Brain Research*, *1363*, 117–127. https://doi.org/10.1016/j.brainres.2010.09.094

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–268. https://doi.org/10.1016/j.tics.2012.04.001

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

Gerster, M., Waterstraat, G., Litvak, V., Lehnertz, K., Schnitzler, A., Florin, E., Curio, G., & Nikulin, V. (2022). Separating neural oscillations from aperiodic 1/f activity: Challenges and recommendations. *Neuroinformatics*, *20*(4), 991–1012. https://doi.org/10.1007/s12021-022-09581-8

Goldberg, H., Preminger, S., & Malach, R. (2014). The emotion–action link? Naturalistic emotional stimuli preferentially activate the human dorsal visual stream. *NeuroImage*, *84*, 254–264. https://doi.org/10.1016/j.neuroimage.2013.08.032

Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The neural bases of emotion regulation: Reappraisal and suppression of negative emotion. *Biological Psychiatry*, *63*(6), 577–586. https://doi.org/10.1016/j.biopsych.2007.05.031

Golkar, A., Lonsdorf, T. B., Olsson, A., Lindstrom, K. M., Berrebi, J., Fransson, P., Schalling, M., Ingvar, M., & Öhman, A. (2012). Distinct contributions of the dorsolateral prefrontal and orbitofrontal cortex during emotion regulation. *PLOS ONE*, *7*(11), 1–9. https://doi.org/10.1371/journal.pone.0048107

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. S. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267. https://doi.org/10.3389/fnins.2013.00267

Grandjean, D. (2021). Brain networks of emotional prosody processing. *Emotion Review*, *13*(1), 34–43. https://doi.org/10.1177/1754073919898522

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*(3), 271–299. https://doi.org/10.1037/1089-2680.2.3.271

Gross, J., Kujala, J., Hämäläinen, M. S., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 694–699. https://doi.org/10.1073/pnas.98.2.694

Grosse-Wentrup, M., & Buss, M. (2008). Multiclass common spatial patterns and information theoretic feature extraction. *IEEE Transactions on Biomedical Engineering*, *55*(8), 1991–2000. https://doi.org/10.1109/TBME.2008.921154

Gyurkovics, M., Clements, G. M., Low, K. A., Fabiani, M., & Gratton, G. (2021a). The impact of 1/f activity and baseline correction on the results and interpretation of time-frequency analyses of EEG/MEG data: A cautionary tale. *NeuroImage*, *237*, 118192. https://doi.org/10.1016/j.neuroimage.2021.118192

Haghani, M., Bliemer, M. C., Farooq, B., Kim, I., Li, Z., Oh, C., Shahhoseini, Z., & MacDougall, H. (2021). Applications of brain imaging methods in driving behaviour research. *Accident Analysis and Prevention*, *154*, 106093. https://doi.org/10.1016/j.aap.2021.106093

Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science*, *366*(6461), 55–58. https://doi.org/10.1126/science.aax0289

Hämäläinen, M. S., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography – Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*(2), 413–497. https://doi.org/10.1103/RevModPhys.65.413

Hamilton, L. S., & Huth, A. G. (2020). The revolution will not be controlled: Natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience*, *35*(5), 573–582. https://doi.org/10.1080/23273798.2018.1499946

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183, Vol. 52). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, *87*, 96–110. https://doi.org/10.1016/j.neuroimage.2013.10.067

Hauk, O. (2004). Keep it simple: A case for using classical minimum norm estimation in the analysis of EEG and MEG data. *NeuroImage*, *21*(4), 1612–1621. https://doi.org/10.1016/j.neuroimage.2003.12.018

He, B. J., Zempel, J. M., Snyder, A. Z., & Raichle, M. E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron*, *66*(3), 353–369. https://doi.org/10.1016/j.neuron.2010.04.020

Herring, J. D., Esterer, S., Marshall, T. R., Jensen, O., & Bergmann, T. O. (2019). Low-frequency alternating current stimulation rhythmically suppresses gamma-band oscillations and impairs perceptual performance. *NeuroImage*, *184*, 440–449. https://doi.org/10.1016/j.neuroimage.2018.09.047

Herrmann, C. S., Fründ, I., & Lenz, D. (2010). Human gamma-band activity: A review on cognitive and behavioral correlates and network models. *Neuroscience & Biobehavioral Reviews*, *34*(7), 981–992. https://doi.org/10.1016/j.neubiorev.2009.09.001

Hipp, J. F., & Siegel, M. (2013). Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. *Frontiers in Human Neuroscience*, *7*, 338. https://doi.org/10.3389/fnhum.2013.00338

Holdgraf, C. R., Rieger, J. W., Micheli, C., Martin, S., Knight, R. T., & Theunissen, F. E. (2017). Encoding and decoding models in cognitive electrophysiology. *Frontiers in Systems Neuroscience*, *11*, 61. https://doi.org/10.3389/fnsys.2017.00061

Hooker, C. I., & Knight, R. T. (2006). The role of lateral orbitofrontal cortex in the inhibitory control of emotion. In *The orbitofrontal cortex*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198565741.003.0012

Houck, J. M., & Claus, E. D. (2020). A comparison of automated and manual co-registration for magnetoencephalography. *PLOS ONE*, *15*(4), e0232100. https://doi.org/10.1371/journal.pone.0232100

Hovsepyan, S., Olasagasti, I., & Giraud, A.-L. (2023). Rhythmic modulation of prediction errors: A top-down gating role for the beta-range in speech processing. *PLOS Computational Biology*, *19*(11), e1011595. https://doi.org/10.1371/journal.pcbi.1011595

Jensen, O., Kaiser, J., & Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in Neurosciences*, *30*(7), 317–324. https://doi.org/10.1016/j.tins.2007.05.001

Jolly, E. (2018). Pymer4: Connecting R and Python for linear mixed modeling. *Journal of Open Source Software*, *3*(31), 862. https://doi.org/10.21105/joss.00862

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*, *89*(1), 221–234. https://doi.org/10.1016/j.neuron.2015.11.028

Joundi, R. A., Jenkinson, N., Brittain, J.-S., Aziz, T. Z., & Brown, P. (2012). Driving oscillatory activity in the human cortex enhances motor performance. *Current Biology*, *22*(5), 403–407. https://doi.org/10.1016/j.cub.2012.01.024

Kang, J.-H., Jeong, J. W., Kim, H. T., Kim, S. H., & Kim, S.-P. (2014). Representation of cognitive reappraisal goals in frontal gamma oscillations. *PLOS ONE*, *9*(11), 1–10. https://doi.org/10.1371/journal.pone.0113375

Khalighinejad, B., Cruzatto da Silva, G., & Mesgarani, N. (2017). Dynamic encoding of acoustic features in neural responses to continuous speech. *Journal of Neuroscience*, *37*(8), 2176–2185. https://doi.org/10.1523/JNEUROSCI.2383-16.2017

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302*(5648), 1181–1185. https://doi.org/10.1126/science.1088545

Kret, M. E., & Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. *Behavior Research Methods*, *51*(3), 1336–1342. https://doi.org/10.3758/s13428-018-1075-y

Krueger, H.-P., Grein, M., Kaussner, A., & Mark, C. (2005). SILAB – A task-oriented driving simulation [Retrieved January 10, 2025]. *Proceedings of the Driving Simulation Conference North America 2005 (DSC-NA 2005)*, 1–9. https://www.nads-sc.uiowa.edu/dscna/2005/papers/SILAB_A_task_oriented_driving_simulation.pdf

Landesanstalt für Medien NRW. (2021, August). *On track: Ergebnisse einer repräsentativen Nutzerbefragung* (Survey report; retrieved January 10, 2025). Landesanstalt für Medien NRW. Düsseldorf, Germany. https://www.medienanstalt-nrw.de/fileadmin/user_upload/NeueWebsite_0120/Themen/Radio/210831_LFM_NRW_On-Track_Nutzerstudie_A4_Print-at-home.pdf

Lee, T.-W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, *11*(2), 417–441. https://doi.org/10.1162/089976699300016719

Lin, F.-H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*, *31*(1), 160–171. https://doi.org/10.1016/j.neuroimage.2005.11.054

Lingelbach, K., Vukelić, M., & Rieger, J. W. (2024). GAUDIE: Development, validation, and exploration of a naturalistic German auditory emotional database. *Behavior Research Methods*, *56*, 2049–2063. https://doi.org/10.3758/s13428-023-02135-z

Lingelbach, K., Gado, S., Wirzberger, M., & Vukelić, M. (2023). Workload-dependent hemispheric asymmetries during the emotion-cognition interaction: A close-to-naturalistic fNIRS study. *Frontiers in Neuroergonomics*, *4*. https://doi.org/10.3389/fnrgo.2023.1273810

Lingelbach, K., & Rieger, J. W. (2025). Neurophysiological basis of emotional face perception and working memory load in a dual-task MEG study. *Human Brain Mapping*, *46*(8), 1–25. https://doi.org/10.1002/hbm.70242

Lu, R., Dermody, N., Duncan, J., & Woolgar, A. (2024). Aperiodic and oscillatory systems underpinning human domain-general cognition. *Communications Biology*, *7*, 1643. https://doi.org/10.1038/s42003-024-07397-7

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marsicano, G., Bertini, C., & Ronconi, L. (2024). Decoding cognition in neurodevelopmental, psychiatric and neurological conditions with multivariate pattern analysis of EEG data. *Neuroscience & Biobehavioral Reviews*, *164*, 105795. https://doi.org/10.1016/j.neubiorev.2024.105795

Mathôt, S., Fabius, J., van Heusden, E., & van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods*, *50*(1), 94–106. https://doi.org/10.3758/s13428-017-1007-2

Michalke, L., Dreyer, A. M., Borst, J. P., & Rieger, J. W. (2023). Inter-individual single-trial classification of MEG data using M-CCA. *NeuroImage*, *273*, 120079. https://doi.org/10.1016/j.neuroimage.2023.120079

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miller, K. J., Sorensen, L. B., Ojemann, J. G., & den Nijs, M. (2009). Power-law scaling in the brain surface electric potential. *PLoS Computational Biology*, *5*(12), e1000609. https://doi.org/10.1371/journal.pcbi.1000609

Morawetz, C., & Basten, U. (2024). Neural underpinnings of individual differences in emotion regulation: A systematic review. *Neuroscience & Biobehavioral Reviews*, *162*, 105727. https://doi.org/10.1016/j.neubiorev.2024.105727

Mosher, J. C., Leahy, R. M., & Lewis, P. S. (1999). EEG and MEG: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, *46*(3), 245–259. https://doi.org/10.1109/10.748978

Mothes-Lasch, M., Miltner, W. H., & Straube, T. (2012). Processing of angry voices is modulated by visual load. *NeuroImage*, *63*(1), 485–490. https://doi.org/10.1016/j.neuroimage.2012.07.005

Navarro, J., Reynaud, E., & Osiurak, F. (2018). Neuroergonomics of car driving: A critical meta-analysis of neuroimaging data on the human brain behind the wheel. *Neuroscience & Biobehavioral Reviews*, *95*, 464–479. https://doi.org/10.1016/j.neubiorev.2018.10.016

Nowosielski, R. J., Trick, L. M., & Toxopeus, R. (2018). Good distractions: Testing the effects of listening to an audiobook on driving performance in simple and complex road environments. *Accident Analysis and Prevention*, *111*, 202–209. https://doi.org/10.1016/j.aap.2017.11.033

Nunes, L., & Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(2), 133–144. https://doi.org/10.1016/S1369-8478(02)00012-8

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: A synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, *1251*(1), 1–24. https://doi.org/10.1111/j.1749-6632.2012.06751.x

Oosterwijk, S., Mackey, S., Wilson-Mendenhall, C., Winkielman, P., & Paulus, M. P. (2015). Concepts in context: Processing mental state concepts with internal or external focus involves different neural systems. *Social Neuroscience*, *10*(3), 294–307. https://doi.org/10.1080/17470919.2014.998840

Ou, J., & Law, S.-P. (2019). Top-down and bottom-up mechanisms as reflected by beta and gamma oscillations in speech perception: An individual-difference approach. *Brain and Language*, *199*, 104700. https://doi.org/10.1016/j.bandl.2019.104700

Palmiero, M., Piccardi, L., Boccia, M., Baralla, F., Cordellieri, P., Sgalla, R., Guidoni, U., & Giannini, A. M. (2019). Neural correlates of simulated driving while performing a secondary task: A review. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01045

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological Psychiatry*, *54*(5), 504–514. https://doi.org/10.1016/S0006-3223(03)00168-9

Pinheiro, A. P., Barros, C., Vasconcelos, M., Obermeier, C., & Kotz, S. A. (2017). Is laughter a better vocal change detector than a growl? *Cortex*, *92*, 233–248. https://doi.org/10.1016/j.cortex.2017.03.018

Popov, T., Steffen, A., Weisz, N., Miller, G. A., & Rockstroh, B. (2012). Cross-frequency dynamics of neuromagnetic oscillatory activity: Two mechanisms of emotion regulation. *Psychophysiology*, *49*(12), 1545–1557. https://doi.org/10.1111/j.1469-8986.2012.01484.x

Quirk, G. J., & Beer, J. S. (2006). Prefrontal involvement in the regulation of emotion: Convergence of rat and human studies. *Current Opinion in Neurobiology*, *16*(6), 723–727. https://doi.org/10.1016/j.conb.2006.07.004

Romeo, Z., Fusina, F., Semenzato, L., Bonato, M., Angrilli, A., & Spironelli, C. (2022). Comparison of slides and video clips as different methods for inducing emotions: An electroencephalographic alpha modulation study. *Frontiers in Human Neuroscience*, *16*, 901422. https://doi.org/10.3389/fnhum.2022.901422

Rule, R. R., Shimamura, A. P., & Knight, R. T. (2002). Orbitofrontal cortex and dynamic filtering of emotional stimuli. *Cognitive, Affective and Behavioral Neuroscience*, *2*(3), 264–270. https://doi.org/10.3758/CABN.2.3.264

Sakihara, K., Hirata, M., Ebe, K., Kimura, K., Ryu, S. Y., Kono, Y., Muto, N., Yoshioka, M., Yoshimine, T., & Yorifuji, S. (2014). Cerebral oscillatory activity during simulated driving using MEG. *Frontiers in Human Neuroscience*, *8*, 975. https://doi.org/10.3389/fnhum.2014.00975

Sasai, S., Boly, M., Mensen, A., & Tononi, G. (2016). Functional split brain in a driving/listening paradigm. *Proceedings of the National Academy of Sciences*, *113*(50), 14444–14449. https://doi.org/10.1073/pnas.1613200113

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Scheunemann, J., Unni, A., Ihme, K., Jipp, M., & Rieger, J. W. (2019). Demonstrating brain-level interactions between visuospatial attentional demands and working memory load while driving using functional near-infrared spectroscopy. *Frontiers in Human Neuroscience*, *12*, 542. https://doi.org/10.3389/fnhum.2018.00542

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009

Schlaffke, L., Lissek, S., Lenz, M., Juckel, G., Schultz, T., Tegenthoff, M., Schmidt-Wilcke, T., & Brüne, M. (2015). Shared and nonshared neural networks of cognitive and affective theory-of-mind: A neuroimaging study using cartoon picture stories. *Human Brain Mapping*, *36*(1), 29–39. https://doi.org/10.1002/hbm.22610

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2021). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, *147*(3), 293. https://doi.org/10.1037/bul0000303

Schweizer, S., Satpute, A. B., Atzil, S., Field, A. P., Hitchcock, C., Black, M., Barrett, L. F., & Dalgleish, T. (2019). The impact of affective information on working memory: A pair of meta-analytic reviews of behavioral and neuroimaging evidence. *Psychological Bulletin*, *145*(6), 566–609. https://doi.org/10.1037/bul0000193

Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*(3), 1060–1075. https://doi.org/10.1016/j.neuroimage.2004.03.032

Sendhilnathan, N., Basu, D., & Murthy, A. (2017). Simultaneous analysis of the LFP and spiking activity reveals essential components of a visuomotor transformation in the frontal eye field. *Proceedings of the National Academy of Sciences*, *114*(24), 6370–6375. https://doi.org/10.1073/pnas.1703809114

Shi, W., Meisner, O. C., Blackmore, S., Jadi, M. P., Nandy, A. S., & Chang, S. W. C. (2023). The orbitofrontal cortex: A goal-directed cognitive map framework for social and non-social behaviors. *Neurobiology of Learning and Memory*, *203*, 107793. https://doi.org/10.1016/j.nlm.2023.107793

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends in Cognitive Sciences*, *23*(8), 699–714. https://doi.org/10.1016/j.tics.2019.05.004

Sonkusare, S., Nguyen, V. T., Moran, R., van der Meer, J., Ren, Y., Koussis, N., Dionisio, S., Breakspear, M., & Guo, C. (2020). Intracranial-EEG evidence for medial temporal pole driving amygdala activity induced by multi-modal emotional stimuli. *Cortex*, *130*, 32–48. https://doi.org/10.1016/j.cortex.2020.05.018

Sonkusare, S., Qiong, D., Zhao, Y., Liu, W., Chen, H., Zhang, L., Li, M., Yang, K., & Wang, R. (2023). Frequency dependent emotion differentiation and directional coupling in amygdala, orbitofrontal and medial prefrontal cortex network with intracranial recordings. *Molecular Psychiatry*, *28*(4), 1636–1646. https://doi.org/10.1038/s41380-022-01883-2

Sonnleitner, A., Simon, M., Kincses, W. E., Buchner, A., & Schrauf, M. (2012). Alpha spindles as neurophysiological correlates indicating attentional shift in a simulated driving task. *International Journal of Psychophysiology*, *83*(1), 110–118. https://doi.org/10.1016/j.ijpsycho.2011.10.013

Spiers, H. J., & Maguire, E. A. (2007). Neural substrates of driving behaviour. *NeuroImage*, *36*(1), 245–255. https://doi.org/10.1016/j.neuroimage.2007.02.032

Srinivasan, R., Winter, W. R., & Nunez, P. L. (2006). Source analysis of EEG oscillations using high-resolution EEG and MEG. *Progress in Brain Research*, *159*, 29–42. https://doi.org/10.1016/S0079-6123(06)59003-X

Taulu, S., & Simola, J. (2006). Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Physics in Medicine and Biology*, *51*(7), 1759–1768. https://doi.org/10.1088/0031-9155/51/7/008

Taulu, S., Simola, J., & Kajola, M. (2005). Applications of the signal space separation method. *IEEE Transactions on Signal Processing*, *53*(9), 3359–3372. https://doi.org/10.1109/TSP.2005.853302

Tavares, T. P., Logie, K., & Mitchell, D. G. (2016). Opposing effects of perceptual versus working memory load on emotional distraction. *Experimental Brain Research*, *234*, 2945–2956. https://doi.org/10.1007/s00221-016-4697-2

Thuwal, K., Banerjee, A., & Roy, D. (2021). Aperiodic and periodic components of ongoing oscillatory brain dynamics link distinct functional aspects of cognition across adult lifespan. *eNeuro*, *8*(5). https://doi.org/10.1523/ENEURO.0224-21.2021

Tucker, D., Luu, P., & Pribram, K. (1995). Social and emotional self-regulation. *Annals of the New York Academy of Sciences*, *769*, 213–239. https://doi.org/10.1111/j.1749-6632.1995.tb38141.x

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review*, *25*(6), 2005–2015. https://doi.org/10.3758/s13423-018-1432-y

van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*(3), 829–858. https://doi.org/10.1002/hbm.20547

van Pelt, S., Heil, L., Kwisthout, J., Ondobaka, S., van Rooij, I., & Bekkering, H. (2016). Beta- and gamma-band activity reflect predictive coding in the processing of causal events. *Social Cognitive and Affective Neuroscience*, *11*(6), 973–980. https://doi.org/10.1093/scan/nsw017

van Veen, B. D., van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Bio-Medical Engineering*, *44*(9), 867–880. https://doi.org/10.1109/10.623056

van Vliet, M., Liljeström, M., Aro, S., Salmelin, R., & Kujala, J. (2018). Analysis of functional connectivity and oscillatory power using DICS: From raw MEG data to group-level statistics in Python. *Frontiers in Neuroscience*, *12*, 586. https://doi.org/10.3389/fnins.2018.00586

Vinck, M., Batista-Brito, R., Knoblich, U., & Cardin, J. A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron*, *86*(3), 740–754. https://doi.org/10.1016/j.neuron.2015.03.028

Vogels, J., Demberg, V., & Kray, J. (2018). The index of cognitive activity as a measure of cognitive processing load in dual task settings. *Frontiers in Psychology*, *9*, 2276. https://doi.org/10.3389/fpsyg.2018.02276

Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, *90*(3), 1195–1268. https://doi.org/10.1152/physrev.00035.2008

Wascher, E., Arnau, S., Gutberlet, I., Karthaus, M., & Getzmann, S. (2018). Evaluating pro- and re-active driving behavior by means of the EEG. *Frontiers in Human Neuroscience*, *12*. https://doi.org/10.3389/fnhum.2018.00205

Whalen, C., Maclin, E. L., Fabiani, M., & Gratton, G. (2008). Validation of a method for coregistering scalp recording locations with 3D structural MR images. *Human Brain Mapping*, *29*(11), 1288–1301. https://doi.org/10.1002/hbm.20465

Wickens, C. D. (2014). The structure of attentional resources. In *Attention and Performance VIII* (pp. 239–257). Psychology Press.

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research*, *156*, 249–268. https://doi.org/10.1016/S0079-6123(06)56013-3

Witteman, J., van Heuven, V. J., & Schiller, N. O. (2012). Hearing feelings: A quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. *Neuropsychologia*, *50*(12), 2752–2763. https://doi.org/10.1016/j.neuropsychologia.2012.07.026

Zhang, C., Stock, A.-K., Mückschel, M., Hommel, B., & Beste, C. (2023). Aperiodic neural activity reflects metacontrol. *Cerebral Cortex*, *33*(12), 7941–7951. https://doi.org/10.1093/cercor/bhad089

5 General Discussion

This thesis advances the understanding of the neurophysiological basis of emotion-cognition interaction by identifying brain regions and neuronal signatures involved in integrating emotional and cognitive processes within dual-task paradigms, where emotional stimuli can be task-relevant and are presented concurrently with a cognitive task. It sheds light on circumstances in which the valence of naturalistic emotional stimuli (facial expressions and speech) and the current level of cognitive load co-modulate information processing and regulatory control.

This work addresses key gaps in the literature by incorporating both positive and negative emotional stimuli, varying levels of cognitive load, and dual-task paradigms in which information conveyed by emotional stimuli becomes temporarily task-relevant. Previous research has predominantly focused on emotionally negative stimuli and a fixed cognitive load level, and has treated emotional stimuli either exclusively as distractions or consistently as task-relevant (see Brockhoff et al., 2022; Dolcos et al., 2011; Schweizer et al., 2019, for reviews).

Two studies combining MEG and eye-tracking addressed the research questions of (1) how emotional and cognitive processes are integrated in dual-task paradigms, and (2) when emotional valence and the current cognitive load level jointly influence information processing and regulatory control. Study I temporally interleaved a facial expression discrimination task with a visuo-spatial WM task (sequential emotion-cognition dual task). Study III employed a near-naturalistic experimental paradigm involving simulated driving while concurrently listening to speech (simultaneous emotion-cognition dual task). Due to the lack of a validated German database of naturalistic emotional speech, Study II created and validated a novel auditory speech database for emotion induction. Using simulation, bootstrapping, and clustering, it evaluated the extent to which the speech sequences met quality criteria for differentiation, salience, and strength, as well as inter-individual generalisability at the behavioural level. The database is publicly accessible and comprises one and a half hours of emotionally negative, positive, and neutral speech sequences. Each sequence is accompanied by detailed, openly available annotations, including second-by-second temporal dynamics of perceived valence and arousal, as well as ratings of discrete emotions, emotional intensity, dominance, and familiarity. The validated stimuli served as experimental material in Study III.

Both MEG–eye-tracking studies (Study I and III) identified a central role of fronto-temporal brain regions in integrating emotional and cognitive processes and in regulating emotional responses to prevent interference. Furthermore, converging evidence from Study I and Study III suggests that emotion-specific processing occurs during perceptual tasks under low, but not high, cognitive load. This was evidenced by reduced effort for positive and increased effort for negative expressions during the discrimination task, as indexed by pupil dilation, and by subjective emotional appraisal of positive speech during driving. However, emotion-specific processing at

low cognitive load did not disrupt task performance in the concurrent cognitive task. In contrast, during the cognitive tasks requiring WM and executive functioning, interference was more likely to occur at high load in response to negative emotional stimuli, possibly due to depleted cognitive resources. To summarise, these results point to a mediating role of task type (i.e., perceptual versus executive elements) in the co-modulation of emotional valence and cognitive load.

The findings offer new insights into the neurophysiological mechanisms that support emotion-cognition integration in socio-emotional contexts and dual tasks. By systematically combining negative, positive, and neutral stimuli with varying cognitive load, this work extends previous knowledge of how valence and load jointly modulate information processing and regulatory control. Crucially, by examining co-modulation during perceptual tasks across varying WM load levels, the thesis revealed that the nature of the task (i.e., whether it engages perceptual or WM processes) determines co-modulatory effects on emotion-specific processing and emotion-based task interference. This finding challenges the distinction between perceptual and WM load proposed in electrophysiological and neuroimaging research (Brockhoff et al., 2022; Lavie, 2005), which underpins, at least in part, the load theory of selective attention and cognitive control (Lavie, 2005; Lavie et al., 2004).

The following sections will integrate the new insights with existing literature and theoretical frameworks on regulatory control (J. J. Gross, 2015b; Ochsner et al., 2012) and cognitive load (Lavie & de Fockert, 2005; Lavie et al., 2004). Figure 5.1 provides a visual summary of the key findings to guide the discussion.

5.1 The Regulatory Fronto-Temporal Axis: Emotion-Generative Processes and Their Integration in Emotion-Cognition Dual Tasks

5.1.1 Emotion-Generative Processes in the Social Brain

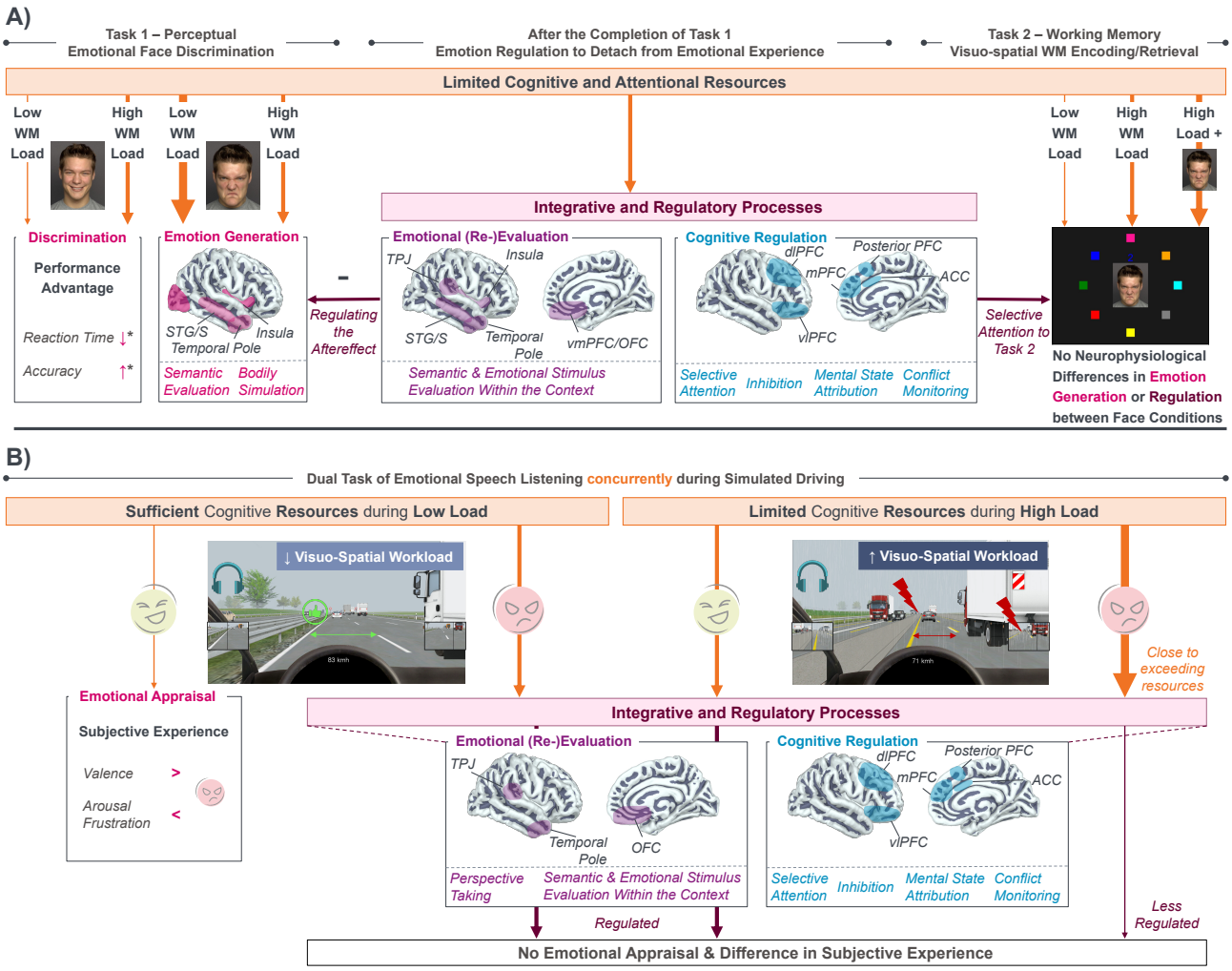
In the facial expression discrimination phase of the sequential emotion-cognition dual task, where emotional information was task-relevant, ERF amplitudes increased in brain regions associated with emotion generation and social cognition in response to negative expressions, regardless of the current workload level. These regions included face-specific occipital areas, inferior and superior temporal regions, the anterior insula, TP, as well as right-lateralised frontal areas, comprising the pPFC, ventrolateral PFC (vlPFC), inferior frontal gyrus (IFG), and vmPFC in the OFC (Figure 5.1A). The occipital and inferior temporal regions of this observed network belong to the visual ventral stream, which is involved in face processing (Bernstein & Yovel, 2015; Fusar-Poli et al., 2009; Vuilleumier & Pourtois, 2007). In line with earlier research, emotional stimuli, and specifically emotionally negative stimuli, increase activation in this network during the perception (Vuilleumier & Pourtois, 2007, see also Section 1.2.1).

In the simultaneous emotion-cognition dual task of driving while listening to speech, beta-band oscillations increased in temporo-parietal regions, including the IPL and TPJ. Based on existing evidence regarding the functional roles of these regions, this increase may reflect enhanced attention, predictive processing, and mentalising in response to emotional speech (J. P. Mitchell, 2009; Ou & Law, 2019; van Pelt

et al., 2016). Additionally, fronto-temporal regions, including the TP and OFC, appear to be involved in the evaluation of emotional speech under varying cognitive load.

Figure 5.1

Visual Summary of the Key Findings from the MEG–eye-tracking Study I (A) and Study III (B)



Note. (A) Sequential dual task with a temporally interleaved emotional face discrimination task (left) and a visuo-spatial n-back task (right). (B) Simultaneous dual task involving simulated driving under low (left) and high (right) workload while listening to emotionally positive (green smiley), neutral, and negative (red smiley) speech. Orange arrow widths represent the assumed cognitive resources allocated to stimulus processing or the cognitive task. Purple arrow widths indicate the extent of the assumed regulatory control exerted. *Discrimination advantages for positive expressions are significant compared to neutral and negative expressions. Illustrated brain regions are adapted from Ochsner et al. (2012), visualising only those regions observed in the two studies and extended to reflect the present findings. WM: working memory; ACC: anterior cingulate cortex; dlPFC: dorsolateral prefrontal cortex; mPFC: medial prefrontal cortex; OFC: orbitofrontal cortex; STG: superior temporal gyrus; STS: superior temporal sulcus; TPJ: temporo-parietal junction; vlPFC: ventrolateral prefrontal cortex; vmPFC: ventromedial prefrontal cortex.

To conclude, both studies identified brain regions implicated in (1) cognitive empathy and ToM (IPL, TPJ, and STG; Adolphs, 2003; Saxe & Kanwisher, 2003; Scott, 2019; and TP; Olson et al., 2007); (2) the internal simulation and evaluation of socio-emotional cues (anterior insula; R. Zhang et al., 2024), as well as (3) evaluation of a stimulus value within the current context (OFC and vmPFC; Adolphs, 2009; Grandjean, 2021). Taken together, the involvement of the outlined brain regions in both studies and their functional interpretation in the literature suggests that the processing of socio-emotional stimuli is intrinsically linked to social cognition and decision-making. Importantly, these regions are also considered key candidates for the regulatory control of emotions (Ochsner et al., 2012, see Figure 5.1, pink-turquoise and pink-purple components).

5.1.2 Regulatory Control in the Integration of Emotion and Cognition

The spatiotemporal dynamics of modulated ERFs during the perception and discrimination of negative, compared to positive, faces indicate a continuous process from emotion generation (i.e., embodied emotion; Niedenthal, 2007) to regulation: At the beginning ($\sim$ 200 ms after the face onset), amplitudes increased in emotion-generating regions (i.e., face-specific occipital areas), followed by increased amplitudes in brain regions categorised by Ochsner et al. (2012) as intermediary between emotion generation and regulation (i.e., the STS, STG, and TP; Figure 5.1A, pink-purple component; $\sim$ 400 to 1,000 ms after the onset). By the end of the discrimination phase ($\sim$ 800 to 1,000 ms after the face onset), increased ERF amplitudes in the ACC, pPFC, and mPFC signalled the recruitment of brain regions categorised as distinctly regulatory (Ochsner et al., 2012, Figure 5.1A, pink-turquoise component). Similarly, in the simultaneous emotion-cognition task, gamma-band modulations in the TP, mPFC, and OFC were implicated in regulating interference effects of emotional speech (Figure 5.1B).

In summary, both studies revealed the involvement of regions that are part of the emotion regulation systems proposed by Ochsner et al. (2012). In the sequential task, these regions were likely recruited to disengage from the emotional experience elicited by negative faces (response-focused regulation through emotional (re-)evaluation starting $\sim$ 400 after the face onset; Figure 5.1A, pink-purple component) and to allocate attentional resources to task-relevant information in the subsequent cognitive task (antecedent-focused regulation through goal-directed attention deployment starting $\sim$ 800 after the face onset; Figure 5.1A, pink-turquoise component; see also J. J. Gross, 2015b). In the simultaneous task, enhanced fronto-temporal gamma oscillations suggested the recruitment of regulatory control. These fronto-temporal gamma oscillations were co-modulated by the valence of emotional speech and the current level of workload, implying that regulatory control is engaged when emotional speech threatens to interfere with driving-related processes and sufficient cognitive resources are available to support this regulation (likely a combination of antecedent- and response-focused regulation; J. J. Gross, 2015b). While the process model of emotion regulation is primarily based on research involving explicit instruction (J. J. Gross, 1998; J. J. Gross & Thompson, 2007; Koole, 2010; Ochsner et al., 2012), the present studies demonstrate that similar neural networks are engaged in regulating emotional responses and integrating emotion and cognition in complex, naturalistic dual-task scenarios, even in the absence of explicit instructions (see also Braunstein et al., 2017; Gyurak et al., 2011; Pruessner et al., 2020).

The following subsections outline a model, synthesised from the literature, of how functional units

within the observed network may specifically contribute to the regulation of emotional interference and the integration of emotional and cognitive processes in dual-task paradigms.

Probing the Emotional Sensation: The Anterior Insula

The insula has been identified as a central hub between emotional and cognitive processes (see R. Zhang et al., 2024, for recent review; see Cauda et al., 2012 and Kurth et al., 2010, for meta-analyses). Study I demonstrated that discriminating negative faces increased activation in the anterior insula and fixation counts on the facial stimulus compared with neutral and positive faces. These findings imply that negative faces are perceived as more salient, supporting the proposed role of the anterior insula as a salience detector for sensory inputs (Menon & Uddin, 2010; Seeley et al., 2007). The insula is functionally organised along the anterior-posterior and dorsal-ventral axes (Cauda et al., 2012; Kurth et al., 2010). In particular, the anterior part of the insula plays a key role in recognising, generating, and regulating emotional responses by probing and re-representing interoceptive sensations (Craig, 2009; Damasio, 2003; Kurth et al., 2010; Lamm & Singer, 2010; Schachter & Singer, 1962; Scherer et al., 2001). Accordingly, the recall of personal emotional episodes increased activity in the insula (Damasio et al., 2000). Furthermore, patients with pronounced anterior insula lesions developed alexithymia, characterised by an impaired awareness of one's own and others' emotional states (Hogeveen et al., 2016).

The dorsal part of the anterior insula is connected to the mPFC and lateral PFC, whereas its ventral part is connected to paralimbic and limbic regions, including the OFC, TP, and amygdala (Lamm & Singer, 2010). The anterior insula is thus also suggested to function as a link between areas involved in mentalising and subcortical emotion-generating regions (Kurth et al., 2010; Lamm & Singer, 2010; R. Zhang et al., 2024).

In the context of the current work, it likely contributes to emotion regulation by probing and modulating the evaluation and intensity of interoceptive emotional sensations. In line with this, the anterior insula is implicated in evaluating the subjective value of a stimulus in light of current goals, thereby guiding decision-making (Damasio, 1996, 2003; Kurth et al., 2010). Previous studies have reported decreased insular activity during the regulation of experienced unfairness in socio-economic games (Grecucci et al., 2013), as well as improved self-regulation and empathy for others' pain after neurofeedback interventions targeting the insula (Yao et al., 2016; see Ruiz et al., 2014, for review).

Labelling Emotions: The Temporal Pole and Ventrolateral Prefrontal Cortex

The TP, the most anterior part of the temporal lobe, is involved in a wide range of cognitive functions, including multisensory integration, face and scene recognition, autobiographical memory, semantic processing, and socio-emotional processing (see Herlin et al., 2021; Olson et al., 2007, for reviews). It is thought to serve as a hub for the convergence of semantic information (Herlin et al., 2021). Olson et al. (2007) further suggested that the TP forms the basis of personal semantic memory and stores linkages between perception and emotion. Like the insula, the TP is tightly connected to paralimbic and limbic regions, including the OFC and amygdala (Olson et al., 2007). The TP may regulate emotions through perceptual and semantic (re-)analysis (see also Goldin et al., 2008; Ochsner et al., 2012; J. Wang et al., 2024).

A second region implicated in emotion regulation through semantic analysis and affective labelling is the vlPFC (see Torre & Lieberman, 2018, for review). Similar to the facial emotion discrimination task in Study I, affective labelling requires participants to provide a descriptive label for the emotional information in a stimulus. Compared to the labelling of non-emotional characteristics (e.g., gender discrimination), affective labelling increases activation in the vlPFC and decreases emotion-generative activity in the amygdala, as well as behavioural and subjective emotional responses (Torre & Lieberman, 2018).

In the studies of this thesis, the TP and vlPFC were involved during the discrimination, labelling, and later regulation (presumed to occur after successful discrimination) of emotionally negative facial expressions (Study I), as well as during the regulation of emotional speech, depending on available cognitive resources (Study II). Ochsner et al. (2012) linked perceptual and semantic analyses during emotion regulation to deliberate, controlled strategies. However, the involvement of the TP and vlPFC in both studies, particularly in Study III, suggests that these processes are, to some extent, also recruited implicitly during emotion-cognition dual tasks, even in the absence of explicit instructions for labelling or regulation.

Self-Other Distinction During Emotional State Inference: The Temporo-Parietal Junction and Medial Prefrontal Cortex

A vital component of ToM and empathy is the ability to disentangle one's own knowledge and feelings from those of others (see Schurz et al., 2021; van Overwalle, 2009, for meta-analyses). The right TPJ facilitates this distinction by shifting attention between the self and others (Krall et al., 2015). Extending from the posterior STS to the IPL, the TPJ is considered an area specific to social cognition (J. P. Mitchell, 2009; Olsson & Ochsner, 2008; Schurz et al., 2021; van Overwalle, 2009). Consistent with the findings of Study III of increased parietal beta- and decreased gamma-band oscillations during emotional speech listening, evidence from electromagnetic and functional brain imaging suggests that beta- and gamma-band oscillations reflect neural mechanisms underlying active inference and predictive listening within the TPJ (Al-Zubaidi et al., 2025; Ou & Law, 2019).

The TPJ is closely connected to the mPFC – another key region involved in mentalising. According to van Overwalle (2009), the TPJ contributes to inferring one's and others' immediate goals and desires, while the mPFC is proposed to integrate long-term knowledge, including self and other traits, as well as socially shared norms and scripts.

The mPFC plays a central role in regulatory control, exhibiting bidirectional connections with the OFC and unidirectional projections to the amygdala (Sonkusare et al., 2023). According to Quirk and Beer (2006), one of its functions in emotion regulation is to maintain the goal of downregulating emotion and transmit this information to the OFC for further evaluation (see also Sonkusare et al., 2023). In addition, the unidirectional projections to the amygdala allow the mPFC to exert direct inhibitory control over emotional responses (Sonkusare et al., 2023). Kang et al. (2014) proposed that gamma oscillations facilitate the bidirectional transfer of regulatory goals and goal-related decisions between the mPFC and OFC. MNE-based source localisation and permutation-based spatial clustering in Study III revealed that emotional valence and current cognitive load jointly modulated gamma-band power in the mPFC and OFC. MVPA further indicated that gamma oscillations

in the OFC play a central role in regulating emotional interference as a function of available cognitive resources. Taken together, these findings corroborate the functional interpretation of Kang et al. (2014) and extend it to emotion-cognition dual-task processing.

Evaluating Emotions in Light of Current Goals: The Orbitofrontal Cortex

The OFC, like the insula, is considered a core region for evaluating the emotional salience and reward value of stimuli in relation to personal goals and the current context (Bechara et al., 2000; Hooker & Knight, 2006; Rolls et al., 2020). While the anterior insula may be more involved in recognising and evaluating others' emotions within this context (Study I), the OFC is proposed to guide dynamic goal-directed decision-making in social contexts through evaluative processes (Bechara et al., 2000; Hooker & Knight, 2006). Given its role in linking emotional information with goal-directed behaviour (Bechara et al., 2000), the OFC is a key candidate region for integrating emotional and cognitive processes. Based on prior functional interpretations and the evidence from this work, it may be concluded that the OFC computes a reward value for emotional information (Rolls et al., 2020) based on the context of the current cognitive demand and available resources during dual tasks. This reward value is then relayed to the mPFC, amygdala, and cingulate cortex (Rolls et al., 2020; Sonkusare et al., 2023). Through its connections to the mPFC, amygdala, and ACC, it likely contributes to the integration of emotional information by signalling a need for regulation (mPFC), guiding emotion generation (amygdala), and facilitating the learning and monitoring of action-outcome relations (ACC; Ochsner et al., 2012; Quirk & Beer, 2006; Rolls et al., 2020).

To summarise, the synthesised model and network of functional units proposes that emotional information is integrated and regulated within cognitive processes by probing the emotional experience (insula and TP; Olsson & Ochsner, 2008), labelling it (TP and vlPFC; Herlin et al., 2021; Torre & Lieberman, 2018), relating it to the self in a social context (TPJ and mPFC; van Overwalle, 2009), and evaluating its value within that context (OFC and insula; Bechara et al., 2000; Kurth et al., 2010).

While the mPFC is suggested to be primarily responsible for regulating activity in subcortical emotion-generating regions (Popov et al., 2012; Quirk & Beer, 2006; Sonkusare et al., 2023), it is likely the OFC and anterior insula that play a pivotal role in integrating emotional information with the current goals of ongoing cognitive processes. In light of the outlined functional interpretation and their involvement in Study I (anterior insula) and Study III (OFC), these two regions emerge as core hubs for integrating emotional and cognitive processes and evaluating the need for regulatory control. These findings offer novel evidence for the neurophysiological basis underlying emotion-cognition integration in dual tasks.

The next chapter outlines how the valence of emotional stimuli and the current level of cognitive load jointly influence the availability of cognitive resources for information processing, as well as integration and regulation of emotional components in emotion-cognition dual tasks. It also examines additional factors that modulate the interaction between emotional valence and cognitive load.

5.2 Under Which Circumstances do Emotional Valence and Cognitive Load Co-Modulate Information Processing and Regulatory Processes?

The findings of both Study I and Study III revealed that, under specific circumstances, emotional valence and cognitive load jointly influence attention allocation, information processing, and regulatory control over emotional information.

In the simultaneous dual task of Study III, joint modulation by valence and workload indicated that emotion processing occurred only for positive speech under low workload (Figure 5.1B). Consistent with this, voluntary attention to and appraisal of positive emotion in speech resulted in a positive subjective experience only under low workload. During high workload, fronto-temporal control mechanisms appeared to down-regulate emotional processing of positive speech; however, their regulatory control over negative speech seemed to diminish, likely due to depleted cognitive resources.

Correspondingly, in the sequential dual task of Study I, gaze-related measures indicated emotion-specific processing, indexed by pupil dilation, during emotion discrimination under WM low load, and showed a trend towards emotional interference, reflected in reduced target-oriented gaze behaviour in response to negative facial distractors during WM encoding and retrieval under high load (Figure 5.1A).

The following subsections discuss these results in light of the load theory of selective attention and cognitive control (Lavie, 2005, 2010; Lavie et al., 2004).

5.2.1 Emotion Processing and Interference Under Various Cognitive Load Levels and Types

A substantial body of research has drawn on Lavie et al. (2004)'s theory to interpret the influence of cognitive load on distractor processing in general, and on the processing of emotional information in distractors specifically (Brockhoff et al., 2022; Murphy et al., 2016). The theory posits a differentiation between perceptual and WM load: It suggests that, for perceptual load, distractor interference and the processing of task-irrelevant information – and, in the case of emotional information, a resulting emotional response – occur at low, but not high, load. Conversely, for WM load, these effects emerge at high, but not low, load (Lavie & de Fockert, 2005; Lavie et al., 2004).

Both Study I and Study III employed emotion-cognition dual tasks that incorporated an emotional component alongside perceptual and WM elements. Their findings advance the understanding of cognitive load effects on temporarily task-relevant emotional stimuli and offer converging novel insights that challenge the proposed differentiation between perceptual and WM load. Specifically, in both studies, emotion-specific processing occurred during low, but not high, load without interfering with the cognitive task (Figure 5.1A and B, left side). In Study I, pupil dilation decreased in response to positive faces and increased for negative faces compared to neutral faces under low cognitive load, reflecting differences in perceptual effort during emotion discrimination (see also van Dillen & Derks, 2012; van Dillen & Koole, 2009). Likewise, subjective ratings in Study III revealed positive appraisal, indicated by higher valence and lower arousal and frustration ratings, during emotionally positive speech compared with neutral or negative speech under low load, but not under high load.

In their review, Murphy et al. (2016) highlighted that the common practice of manipulating cognitive load through perceptual tasks (see Lavie, 2005, for review) introduces a risk of circular reasoning. By disentangling the cognitive load manipulation from the perceptual task, the present results suggest that the previously proposed differentiation by load type may instead be determined by the nature of the task rather than the source of cognitive load. Emotion-specific processing appears to occur under low WM load or workload in perceptual tasks (Study I: perceptual processing difference indexed by pupil dilation; Study III: perceptual subjective experience). In contrast, the combination of emotionally negative stimuli and high WM load or workload seems to elicit interference with the cognitive task and goal-directed regulatory processes (Study I: reduced goal-directed gaze behaviour; Study III: reduced emotion regulation on a neurophysiological level; Figure 5.1A and B, right side).

5.2.2 Emotional Processing and Cognitive Load in Naturalistic Paradigms

While both studies employed naturalistic emotional stimuli, the degree of naturalism in the cognitive task components varied. Study I used a classical, abstract WM task (see Figure 2.3), while Study III involved a highly realistic driving simulation (see Figure 4.1). Although the driving task incorporated a substantial proportion of perceptual components, the observed emotion-specific interference during high cognitive load is consistent with effects proposed for WM load but not perceptual load paradigms (Lavie & de Fockert, 2005; Lavie et al., 2004). The current evidence raises the question of whether, and how frequently, purely perceptual tasks – which underpin much of the evidence for the load theory (Lavie, 2010) – are encountered in realistic environments.

The present results further revealed considerable differences in the processing of naturalistic emotional stimuli compared to that of isolated, artificial ones. In Study III, participants appeared to inhibit the emotional impact of continuous negative speech already during low cognitive load, as evidenced by increased fronto-temporal gamma oscillations and no differences in subjective ratings between negative and neutral speech. The results indicated that regulatory control via fronto-temporal gamma oscillations was effective under low cognitive load but significantly diminished for negative speech under high load, as evidenced by reduced fronto-temporal gamma-band power (see also Lingelbach et al., 2023, for similar results with functional near-infrared spectroscopy, fNIRS).

The findings in Study III differ from those of studies with similar research questions that employed an event-related design using context-free single words as distractors and a commonly applied, abstract perceptual task (Mothes-Lasch et al., 2012). We observed reduced lateralisation and greater involvement of higher-order cortical areas during the processing of naturalistic emotional speech compared to studies using isolated emotional single words (see Hamilton & Huth, 2020, for a review on naturalistic speech). In their fMRI study, Mothes-Lasch et al. (2012) observed increased activation in the right superior temporal region during low, but not high, perceptual load. In contrast to findings from Study III, those authors concluded that negative emotional words are attended to and induce interference effects during low, but not high, load (Mothes-Lasch et al., 2012).

During the processing of photorealistic faces in Study I, we observed that participants applied specific

template-matching strategies (Leopold et al., 2001, 2006) to discriminate positive expressions from other emotions based on low-level features (e.g., visible teeth). Such processing strategies appear to be specific to photorealistic faces and are less observed for artificially created and cut-out grey-scaled faces (see discussion in the review by Xu et al., 2021). Notably, artificial and cut-out grey-scaled faces are employed in at least a third of studies investigating the effects of facial emotional stimuli on selective attention and cognitive control (see Schindler & Bublatzky, 2020, for review). This may explain why the findings in Study I, which showed increased ERF amplitudes for emotionally negative expressions, partially diverged from prior work. Previous studies using more artificial facial stimuli reported enhanced evoked responses to both emotionally negative and positive faces (see Schindler & Bublatzky, 2020, Table 2). Importantly, this suggests that earlier findings may have limited generalisability to naturalistic contexts, thereby highlighting the contribution and relevance of Study I.

Taken together, the findings from both studies indicate that cognitive tasks and emotional stimuli are processed differently in naturalistic compared with artificial contexts. Naturalistic experimental paradigms, which closely approximate real-world phenomena, are considered to enhance the generalisability of findings because they involve less abstraction than artificially isolated stimuli and tasks (Fan et al., 2021; Hamilton & Huth, 2020; Holleman et al., 2020). The extent to which research insights translate to real-world settings is referred to as ecological validity (Holleman et al., 2020). In conclusion, a key finding of this thesis is that a substantial degree of naturalism is essential to ensure generalisability when investigating the interaction of emotional and cognitive processes.

5.2.3 Emotional Interference in Emotion-Cognition Dual Tasks

The second research question in this thesis further examined the circumstances under which emotional interference arises from the interaction of emotional valence and cognitive load. Both studies revealed interference tendencies by emotionally negative but not positive stimuli during high cognitive load. In Study I, negative distractors tended to reduce goal-directed gaze behaviour under high WM. In the simultaneous emotion-cognition dual task of Study III, regulatory control over negative speech diminished during high visuo-spatial workload. Three key conclusions can be drawn from these observations.

Firstly, the source of interference likely differed between the studies. In Study I, interference can be attributed to a combined effect of selective attention (Pashler, 1994; Salvucci & Taatgen, 2008) and capacity limits in attentional control (Cowan, 1999; Wickens, 2002). In Study III, interference is primarily explained by capacity limits (Cowan, 1999) or by concurrent processing demands on central cognitive resources (Salvucci & Taatgen, 2008), as the emotional and visuo-spatial cognitive information was presented through different sensory modalities. This engaged distinct perceptual resource channels and reduced the likelihood of a selective attention bottleneck (Salvucci & Taatgen, 2008; Wickens, 2002). Notably, the bottleneck effect observed in Study I, reflected in reduced attention to task-relevant information presented within the same sensory modality as the emotional distractors, may be more critical in terms of performance impairment in cognitive tasks (Norman & Shallice, 1986).

Secondly, the temporal proximity of the emotional and cognitive task components seems to be a critical factor in determining whether interference occurs and whether it manifests at the neural level. In Study I,

during the sequential dual task, interference was observed in gaze-related measures, without affecting neural processing. However, in the simultaneous dual task of Study III, interference affected neural regulatory control mechanisms. It appears that the temporal window between the sequential task components of Study I was sufficiently long to circumvent neural interference (Duncan et al., 2021; Pashler, 1994).

Thirdly, both studies identified distinct compensatory and regulatory mechanisms that prevented severe performance interference. To maintain performance under high cognitive load, compensatory stimulus-driven processing increased, as indicated by decreased temporo-occipital alpha oscillations during WM encoding (Study I), and by increased gamma- and decreased beta-band power associated with spatial and sensorimotor processing during driving (Study III). Regulatory processes in cognitive control regions appeared to contribute to an effective switch from the emotional to the cognitive task for negative faces at the end of the discrimination phase (Figure 5.1A), while gamma oscillations in the OFC seemed to support interference regulation in the simultaneous emotion-cognition dual task (Figure 5.1B). Importantly, the emotional valence of stimuli influenced the activation of regulatory processes in both studies. However, joint modulation by valence and cognitive load on regulatory control occurred only during the simultaneous dual task.

5.3 Limitations and Ongoing Challenges

While this thesis provides important insights into the neural integration of emotional and cognitive processes, as well as the conditions under which emotional valence and cognitive load interact in emotion-cognition dual-task paradigms, certain limitations must be acknowledged. These pertain to the intensity of emotional stimuli and their ability to induce emotional responses, the characteristics of emotion-cognition dual-task paradigms, and limitations of interpreting divergent evidence in a multimodal measurement approach.

5.3.1 Effects of Stimulus Characteristics

The studies aimed to induce emotional processing by presenting socio-emotional, naturalistic stimuli (i.e., emotional facial images and speech sequences) within dual-task paradigms. However, particularly in Study I, emotional interference under high cognitive load was modest. Moreover, evidence for elicited emotional responses and interference did not converge across subjective, behavioural, and neurophysiological measures in either study.

The emotional intensity of a stimulus and its emotional engagement play a crucial role in reliably inducing emotional responses during both explicit and implicit processing (Fan et al., 2021). Stimulus arousal appears to modulate perceived emotional intensity (García-Pacios, Del Río, et al., 2015) and neurophysiological emotional processing (García-Pacios, Garcés, et al., 2015). Given the limited number of studies directly comparing arousal and valence across stimulus types (but cf. Britton et al., 2006), it is possible that certain stimuli, such as emotional facial images, generally elicit weaker emotional responses than others, such as emotional scenes (Britton et al., 2006; Ochsner & Gross, 2005; but cf. Carretié, 2014). Importantly, this may stem from the static, two-dimensional, and therefore still artificial nature of facial images. Recent studies have shown that three-dimensionality of emotional stimuli enhances perceived arousal, emotional intensity, and

neurophysiological responses (see M. H. Jamil et al., 2021, for facial expressions, and Tian et al., 2021, for emotional scenes).

Regarding the naturalistic speech stimuli, a novel database, GAUDIE, was developed explicitly for emotion induction. The stimuli were annotated for emotional intensity and thoroughly validated to elicit negative, positive, and neutral emotional responses. Despite its behavioural validation, the GAUDIE database has yet to be assessed at the neurophysiological level. Such validation would help elucidate how presenting stimuli within a dual-task paradigm modulates neural emotional responses. Notably, the behavioural validation sample consisted of young, highly educated individuals, reflecting the typical population in neurophysiological studies. This may, however, limit the generalisability of the stimulus annotations to this population. In particular, stimuli designed to elicit positive emotions through humour may be influenced by age-specific, cultural, and contemporary preferences. Accordingly, the stimulus database will require a revalidation over time to maintain its validity.

5.3.2 Effects of Dual-Task Characteristics

The emotion-cognition dual tasks in Study I and Study III differed in three critical aspects: (1) the temporal proximity of the emotional and cognitive tasks (sequential vs. simultaneous), (2) whether the task instructions explicitly directed attention to emotional information (emotion recognition vs. content recall), and (3) the time interval between presenting the emotional stimulus and providing a response to the emotional task (immediate vs. delayed). While the studies provided some insights into the effects of these differences, the interpretability of other findings is limited, as these parameters were not systematically manipulated within each emotion-cognition dual task. Further systematic investigations are required to answer the following questions: What is the critical time interval required for effective integration of emotional and cognitive processes, and at what point does interference emerge? How does the explicit focus on emotional content influence emotional processing and regulation in the emotion-cognition dual tasks (see also Rellecke et al., 2012; Torre & Lieberman, 2018, for discussion)? How does a regularly prompted behavioural response to the emotional task influence information processing and emotional interference in the dual tasks? In particular, computational approaches modelling dual-task mechanisms within a cognitive architecture (e.g., Salvucci & Taatgen, 2008) could shed light on temporal dynamics and sources of task interference (see also Held et al., 2024).

5.3.3 Effects of Measurement Choices

The integration of eye-tracking, behavioural, and subjective measures within the MEG studies enabled a multidimensional investigation of the interaction between emotional and cognitive processes. This approach has the advantage of identifying converging evidence and facilitating the interpretation of findings (Braunstein et al., 2017). However, results observed in one measurement modality but not another can be difficult to interpret. For instance, in Study I, stimulus valence and WM load jointly modulated gaze-related measures, but not behavioural or neurophysiological correlates. Further joint analyses of fixation-related evoked brain responses to negative facial distractors potentially reveal the corresponding neurophysiological basis of the

interaction between emotional valence and load, thereby complementing the interpretation of gaze-related effects (Spiering & Dimigen, 2024).

Finally, although MEG is well-suited for investigating emotional and cognitive interactions given its excellent temporal and adequate spatial resolution, the detection of subcortical modulations with MEG remains a methodological challenge (Hansen et al., 2010; Hämäläinen et al., 1993; but cf. Pizzo et al., 2019). Consequently, techniques such as intracranial electrophysiological recordings (e.g., Weidner et al., 2024) or fMRI (Ochsner et al., 2012) may be better suited for investigating modulations related to emotion generation and its top-down regulation within subcortical structures such as the amygdala.

While the current section addressed specific limitations and ongoing challenges identified in the two research projects, the next section will reflect on how the research field of interacting emotional and cognitive processes may evolve in the future.

5.4 Future Directions of the Research Field

In their review, Fan et al. (2021) introduced a five-level categorisation system, reflecting the degree of naturalism in social neuroscience research. Within this framework, Study I falls under Level 1, due to its use of static facial images as stimuli. Study III spans from Level 2, involving dynamic but pre-recorded auditory speech sequences, to Level 3, reflecting contingent interactions with virtual car agents in the driving simulation. Level 4 encompasses real-life social interaction with other individuals within laboratory settings, whereas Level 5 extends this multi-agent research into naturalistic, real-world environments. Importantly, the increase in naturalism and ecological validity comes at the cost of experimental control (Fan et al., 2021).

Advancing research on the interaction between emotional and cognitive processes will require further well-controlled studies at Level 2 and Level 3 to replicate and extend current findings on the underlying neural mechanisms at the individual brain level. This includes dual-task studies employing multisensory, dynamic emotional stimuli, as well as stimuli that convey subtle emotional nuances encountered in everyday social interactions (e.g. using movie excerpts; Labs et al., 2015; E. Yang et al., 2023; and 3D stimuli; Tian et al., 2021). Study III demonstrated that MVPA decoding holds considerable potential for investigating the interplay between emotion and cognition, as well as for exploring inter-individual differences. Besides MVPA decoding, encoding approaches can further shed light on how specific stimulus feature representations (e.g., acoustic and linguistic features; Desai et al., 2021) or top-down context-based predictions (e.g., based on semantics; Broderick et al., 2019) influence neural information processing in complex multisensory environments.

Additionally, research at higher levels of naturalism (Level 4 and 5) should investigate how multi-agent contexts and real-life social interactions affect the interplay between emotional and cognitive processes (Acconito et al., 2025). In this regard, hyperscanning techniques present a valuable opportunity to study shared neural dynamics of emotional integration and regulation across interacting individuals (Balconi & Vanutelli, 2017; Hasson et al., 2012).

5.5 Practical Implications and Translational Potential

Research into the neurophysiological basis of emotion-cognition interactions offers compelling translational potential for advancing adaptive assistive systems in safety-critical and performance-driven environments, as well as for informing clinical interventions.

Brain-Computer Interfaces (BCIs) play a central role in these applications. They continuously monitor neural activity to decode brain signatures of cognitive and emotional states (passive BCI) or intentional control commands (active BCI), which are then relayed to external systems such as computers, robotic devices, or prosthetic limbs. These systems can provide contingent feedback on the recognised mental states or intentions, and adapt accordingly, forming a closed-loop system. Although BCIs primarily target brain activity, they can also rely on or integrate other physiological and bodily signals in a hybrid approach (Nicolas-Alonso & Gomez-Gil, 2012; Zander & Kothe, 2011).

In this context, the findings of the thesis hold relevance for both healthy and clinical populations to: (1) inform targeted (clinical) interventions, including neurofeedback training and non-invasive neuromodulation techniques (e.g., transcranial electrical or magnetic stimulation), on key neural substrates in the emotion-cognition interaction (e.g., Huang et al., 2023; Ruiz et al., 2013, 2014; Sabharwal et al., 2016), (2) detect state- and trait-level impairments in the emotion-cognition integration and regulatory control over emotional interference (e.g., Sabharwal et al., 2016; see Y. Chen et al., 2025, for a recent review), and (3) develop closed-loop assistive systems that provide customised, context-sensitive support for integrating and regulating emotional processes during cognitive tasks (Y. Chen et al., 2025).

The present work identified the anterior insula and OFC as key regions supporting the integration and regulation of emotional information. This highlights their potential as targets for diagnostic assessment, therapeutic intervention, and informative brain signatures in neuro-adaptive systems. Study III further suggests that gamma oscillations are implicated in transferring information across fronto-temporal regions during regulatory control (see also Kang et al., 2014; Sonkusare et al., 2023). These insights are particularly relevant given the persisting uncertainty over which brain signatures are most suitable for effectively modulating emotion regulation abilities through stimulation (Tanaka et al., 2025).

Supporting evidence for the suitability of these identified brain signatures comes from neurofeedback studies (see Ruiz et al., 2014, for review), as well as from animal research (W. Cao et al., 2018; Yin et al., 2024). A real-time fMRI study reported that longitudinal neurofeedback training enabled schizophrenia patients to voluntarily up- and down-regulate activity in the bilateral anterior insula during facial emotion recognition. This self-regulatory ability improved behavioural accuracy in recognising negative facial expressions and enhanced effective connectivity between the anterior insula, mPFC, and amygdala (Ruiz et al., 2013). An EEG neurofeedback study with healthy human participants emphasised the role of fronto-temporal gamma and beta oscillatory activity for controlling emotional experiences when watching negative videos (Huang et al., 2023).

In addition, animal research in mice points to the transdiagnostic therapeutic potential of modulating frontal gamma oscillations to improve regulatory control. Using an inflammation-triggering molecule, Yin et al. (2024) induced depression-like states in mice and observed reduced gamma oscillations in the mPFC. In contrast, administration of both rapid- and slow-acting antidepressants increased gamma oscillations in the

mPFC (Yin et al., 2024). Additionally, optogenetic stimulation of the mPFC in the gamma range positively modulated social behaviour in mice with an autism-associated mutation (W. Cao et al., 2018). Importantly, these studies examined emotion regulation in isolation, without integrating ongoing cognitive processes or context-dependent variations in cognitive load. Further research should investigate whether the findings on voluntary regulation or causal manipulation of gamma oscillations also extend to the OFC during dual-tasking, thereby facilitating the integration of emotional processes into cognitive functioning and context-dependent regulation.

To conclude, the ability to integrate and regulate emotional information during cognitive functioning has long-term implications for well-being and mental health, and facilitates social interaction (Aldao et al., 2010; J. J. Gross & Jazaieri, 2014). Consequently, it supports the inclusion of clinical populations in occupational settings and is generally essential for optimal performance across various applications, particularly in learning and safety-critical environments.

In educational and training environments, emotional experiences, such as frustration and distress, significantly affect the learning process. The ability to regulate one's emotions, inhibit emotional distractions, and control attention allocation is critical for effective learning and the retrieval of acquired knowledge. Consequently, BCIs and neuro-adaptive tutoring systems that train and support the integration and regulation of emotional states can enhance attentional control, optimise learning strategies, and ultimately improve learning outcomes (see N. Jamil et al., 2021, for review).

By investigating dual-tasking during driving, Study III directly addressed another highly important, safety-relevant application: Distracted driving remains a major cause of accidents, posing risks not only to vehicle occupants but also to other road users (National Center for Statistics and Analysis, 2025, part of the National Highway Traffic Safety Administration, NHTSA). Both the European Commission and the NHTSA highlighted the urgent need to reduce distracted driving. The European Commission specifically called for research aimed at recognising mental states that compromise driving safety and for the development of reliable warning mechanisms (European Commission & Directorate General for Mobility and Transport, 2020).

The results of Study III provide implications for safety guidelines and diagnostic neural and ocular signatures linked to driving under emotional interference and high cognitive load. In particular, gaze-related metrics show great promise due to seamless integration of eye-tracking hardware into vehicle interiors (Murali et al., 2022) and advances in wearable sensor technologies (Z. Wang et al., 2025). These metrics comprised reduced gaze dispersion and increased pupil diameter during increased cognitive load, as well as abrupt fluctuations in pupil diameter, which may reflect attentional control. Across both studies, pupil diameter proved to be a robust index of perceptual effort in visual processing (see van der Wel & van Steenbergen, 2018, for review), but not in auditory processing within a multisensory paradigm (cf. Winn et al., 2015, for auditory load effects on pupil dilation in a purely auditory paradigm). Interestingly, exploratory analyses of abrupt fluctuations in pupil diameter suggested a potential link to regulatory control over auditory emotional interference – a finding that warrants further replication and investigation (see also Demberg & Sayeed, 2016; Vogels et al., 2018). An informative signature pattern combining gaze dispersion, pupil diameter, and its fluctuations could enable differentiation between visual driving-related and auditory distraction-related cognitive load.

Taken together, these insights hold significant potential for adaptive systems beyond the driving context, enabling feedback mechanisms to alert users to distractions in multisensory environments and offer support in challenging situations. To advance this line of research in the context of neuro-adaptive systems, future studies using mobile neurophysiological measurement hardware should evaluate the reliability of fronto-temporal gamma oscillations as a neural marker of emotion-cognition integration during real-life driving.

5.6 Concluding Remarks

The thesis addresses two key research questions and gaps in the existing literature on emotion-cognition interaction. Firstly, how are emotional and cognitive processes integrated in dual-task paradigms where emotional information can be task-relevant and is concurrently presented during a cognitive task? Secondly, under which circumstances do the emotional valence of a stimulus (as conveyed through facial expressions and speech) and cognitive load jointly influence the processing of information and regulatory control?

This work also contributes a novel, validated database of naturalistic emotional speech. It is the first and currently only available German resource suitable for researching the processing and induction of emotions during the perception of meaningful, naturalistic speech sequences. The database comprises 92 minutes of emotionally positive, neutral, and negative speech sequences, and is publicly accessible alongside a comprehensive set of annotations for each sequence. These include second-by-second ratings of perceived valence and arousal, as well as evaluations of discrete emotions, emotional intensity, dominance, and familiarity. Clustering and simulation analyses, combined with bootstrapping, quantified emotion differentiation, salience, strength, and inter-individual generalisability.

Regarding the first research question on the neurophysiological basis of emotion-cognition integration, findings from two MEG–eye-tracking studies indicate that fronto-temporal brain regions support this integration by regulating emotional responses according to available cognitive resources. In line with neuroscientific studies conducted within the framework of the process model of emotion regulation (Ochsner et al., 2012), this regulatory network encompasses frontal regions, such as the vlPFC, vmPFC, mPFC, and ACC, the TP, as well as parietal regions, such as the IPL and TPJ. A novel insight of this thesis is the extension of the network to include the anterior insula and OFC, which appear to play a key role in integrating and regulating emotional information during dual tasks. The anterior insula and OFC are likely responsible for assessing the value of emotional stimuli in relation to the context, demands, and goals of other cognitive processes. In light of previous research on the functional roles of these brain regions, regulatory processes within the identified network are proposed to involve probing and simulating the experience associated with the emotional stimulus, labelling it, distancing from it through self-other distinction (particularly in social contexts), and evaluating its relevance in relation to current task demands and ongoing cognitive processes in the dual-task paradigm.

Regarding the second research question, both studies suggest that emotion-specific processing occurs only under low cognitive load during perceptual tasks, reflecting a joint influence of emotional valence and cognitive load. This emotion-specific processing includes valence-specific responses and emotional

appraisal, without interfering with the cognitive task. Positive emotional stimuli, when paired with low – but not high – cognitive load, reduced pupil dilation and elicited more positive subjective experiences. Notably, discrimination accuracy for positive facial stimuli was enhanced regardless of the cognitive load level. Thus, reduced discrimination effort, but not accuracy, as well as emotional appraisal of positive stimuli, appear to depend on the availability of sufficient cognitive resources. Conversely, high cognitive load elicited emotional interference in response to negative – but not positive – emotional stimuli, likely due to diminished cognitive resources. This was demonstrated by reduced target-directed gaze behaviour in the sequential dual task of Study I and reduced gamma-mediated regulatory processes in the simultaneous dual task of Study III. Enhanced stimulus-driven and compensatory processing modes appeared to mitigate severe task interference under high visuo-spatial load. These modes were characterised by reduced temporo-occipital alpha-band power during WM encoding in Study I, and enhanced spatial cognition and sensorimotor processing through decreased beta and increased gamma oscillations in Study III. The findings substantially extend existing knowledge on how emotional valence and cognitive load co-modulate information processing and regulatory control by integrating emotionally negative, neutral, and positive stimuli across varying levels of cognitive load and employing a naturalistic experimental approach.

The disentanglement of cognitive load manipulation from perceptual task components provides novel evidence that co-modulation and emotion-specific processing are influenced more by the nature of the task (targeting perceptual processing vs. executive functioning and WM) than by the source of cognitive load (i.e., perceptual load vs. WM load). This, in turn, challenges the concept of perceptual load and the distinction between perceptual and WM load effects proposed in earlier electrophysiological and neuroimaging studies (Brockhoff et al., 2022; Lavie, 2005).

Further well-controlled studies are needed to examine how emotion-cognition interactions and associated interferences are affected by specific dual-task parameters, such as varying temporal proximity or task instruction. In parallel, this research field would benefit greatly from moving towards even more naturalistic experimental paradigms to extend our understanding of how the human brain integrates emotion and cognition in its natural environment. This also includes real-life social interactions and shared experiences of interacting emotional and cognitive processes, both inside and outside the laboratory.

Understanding the neurophysiological basis of the emotion-cognition interaction and integration has significant potential for improving adaptive assistive systems in safety-critical and performance-driven environments, and for informing clinical interventions. Altogether, this work lays the groundwork for exciting new directions in both fundamental research and the development of neuro-adaptive technologies that support the dynamic interplay between emotional and cognitive functioning in naturalistic settings.

General References

Acconito, C., Rovelli, K., Saquella, F., & Balconi, M. (2025). Cognitive and emotional engagement in negotiation: Insights from EEG and fNIRS hyperscanning. *Experimental Brain Research*, *243*(6), 151. https://doi.org/10.1007/s00221-025-07093-w

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, *12*(2), 169–177. https://doi.org/10.1016/S0959-4388(02)00301-X

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, *4*(3), 165–178. https://doi.org/10.1038/nrn1056

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, *60*, 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, *30*(2), 217–237. https://doi.org/10.1016/j.cpr.2009.11.004

Almeida, P. R., Ferreira-Santos, F., Chaves, P. L., Paiva, T. O., Barbosa, F., & Marques-Teixeira, J. (2016). Perceived arousal of facial expressions of emotion modulates the N170, regardless of emotional category: Time domain and time–frequency dynamics. *International Journal of Psychophysiology*, *99*, 48–56. https://doi.org/10.1016/j.ijpsycho.2015.11.017

Alpert, M., & Rosen, A. (1990). A semantic analysis of the various ways that the terms “affect,” “emotion,” and “mood” are used. *Journal of Communication Disorders*, *23*(4), 237–246. https://doi.org/10.1016/0021-9924(90)90002-G

Al-Zubaidi, A., Schepers, I. M., Beck, A.-K., Schwabe, K., Runge, J., Abdallat, M., Krauss, J. K., & Rieger, J. W. (2025). Word onset tracking in neural responses of human basal ganglia nuclei. *Brain Structure and Function*, *230*(6), 105. https://doi.org/10.1007/s00429-025-02968-8

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277. https://doi.org/10.1038/nrn1884

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baker, R. S., D’Mello, S. K., Rodrigo, M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners’ cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223–241. https://doi.org/10.1016/j.ijhcs.2009.12.003

Balconi, M., & Vanutelli, M. E. (2017). Cooperation and competition with hyperscanning methods: Review and future application to emotion domain. *Frontiers in Computational Neuroscience*, *11*. https://doi.org/10.3389/fncom.2017.00086

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, *10*(3), 295–307. https://doi.org/10.1093/cercor/10.3.295

Becker, D. V., & Rheem, H. (2020). Searching for a face in the crowd: Pitfalls and unexplored possibilities. *Attention, Perception, and Psychophysics*, *82*(2), 626–636. https://doi.org/10.3758/s13414-020-01975-7

Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *9*(9), 1395–1403. https://doi.org/10.1093/scan/nst124

Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience & Biobehavioral Reviews*, *55*, 536–546. https://doi.org/10.1016/j.neubiorev.2015.06.010

Bonnefond, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Current Biology*, *22*(20), 1969–1974. https://doi.org/10.1016/j.cub.2012.08.029

Borst, J., Taatgen, N., & Rijn, H. (2010). The problem state: A cognitive bottleneck in multitasking. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(2), 363–82. https://doi.org/10.1037/a0018106

Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S. S., Muñoz, L. D., Mullinger, K. J., Tierney, T. M., Bestmann, S., Barnes, G. R., Bowtell, R., & Brookes, M. J. (2018). Moving magnetoencephalography towards real-world applications with a wearable system. *Nature*, *555*(7698), 657–661. https://doi.org/10.1038/nature26147

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Braunstein, L. M., Gross, J. J., & Ochsner, K. N. (2017). Explicit and implicit emotion regulation: A multi-level framework. *Social Cognitive and Affective Neuroscience*, *12*(10), 1545–1557. https://doi.org/10.1093/scan/nsx096

Britton, J. C., Taylor, S. F., Sudheimer, K. D., & Liberzon, I. (2006). Facial expressions and complex IAPS pictures: Common and differential networks. *NeuroImage*, *31*(2), 906–919. https://doi.org/10.1016/j.neuroimage.2005.12.050

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press. https://doi.org/10.1037/10037-000

Brockhoff, L., Schindler, S., Bruchmann, M., & Straube, T. (2022). Effects of perceptual and working memory load on brain responses to task-irrelevant stimuli: Review and implications for future research. *Neuroscience & Biobehavioral Reviews*, *135*, 104580. https://doi.org/10.1016/j.neubiorev.2022.104580

Broderick, M. P., Anderson, A. J., & Lalor, E. C. (2019). Semantic context enhances the early auditory encoding of natural speech. *Journal of Neuroscience*, *39*(38), 7564–7575. https://doi.org/10.1523/JNEUROSCI.0584-19.2019

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

Calvo, M. G., & Beltrán, D. (2013). Recognition advantage of happy faces: Tracing the neurocognitive processes. *Neuropsychologia*, *51*(11), 2051–2061. https://doi.org/10.1016/j.neuropsychologia.2013.07.010

Cao, F., Zeng, K., Li, W., Liu, S., Zhang, L., Katembu, S., & Xu, Q. (2022). Influence of scene-based expectation on facial expression perception: The moderating effect of cognitive load. *Biological Psychology*, *168*, 108247. https://doi.org/10.1016/j.biopsycho.2021.108247

Cao, W., Lin, S., Xia, Q.-Q., Du, Y.-L., Yang, Q., Zhang, M.-Y., Lu, Y.-Q., Xu, J., Duan, S.-M., Xia, J., Feng, G., Xu, J., & Luo, J.-H. (2018). Gamma oscillation dysfunction in mPFC leads to social deficits in neuroligin 3 R451C knockin mice. *Neuron*, *97*(6), 1253–1260. https://doi.org/10.1016/j.neuron.2018.02.001

Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: A review. *Cognitive, Affective and Behavioral Neuroscience*, *14*(4), 1228–1258. https://doi.org/10.3758/s13415-014-0270-2

Cauda, F., Costa, T., Torta, D. M., Sacco, K., D'Agata, F., Duca, S., Geminiani, G., Fox, P. T., & Vercelli, A. (2012). Meta-analytic clustering of the insular cortex: Characterizing the meta-analytic connectivity of the insula when involved in active tasks. *NeuroImage*, *62*(1), 343–355. https://doi.org/10.1016/j.neuroimage.2012.04.012

Chan, M., & Singhal, A. (2015). Emotion matters: Implications for distracted driving. *Safety Science*, *72*, 302–309. https://doi.org/10.1016/j.ssci.2014.10.002

Chen, X., Sorenson, E., & Hwang, K. (2023). Thalamocortical contributions to working memory processes during the n-back task. *Neurobiology of Learning and Memory*, *197*, 107701. https://doi.org/10.1016/j.nlm.2022.107701

Chen, Y., Peng, Y., Tang, J., Camilleri, T., Camilleri, K., Kong, W., & Cichocki, A. (2025). EEG-based affective brain–computer interfaces: Recent advancements and future challenges. *Journal of Neural Engineering*, *22*(3), 031004. https://doi.org/10.1088/1741-2552/ade290

Cohen, D. (1972). Magnetoencephalography: Detection of the brain's electrical activity with a superconducting magnetometer. *Science*, *175*(4022), 664–666. https://doi.org/10.1126/science.175.4022.664

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press.

Craig, A. D. (2009). How do you feel — now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*, 59–70. https://doi.org/10.1038/nrn2555

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1413–1420. https://doi.org/10.1098/rstb.1996.0125

Damasio, A. R. (2003). Feelings of emotion and the self. *Annals of the New York Academy of Sciences*, *1001*(1), 253–261.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, *3*(10), 1049–1056. https://doi.org/10.1038/79871

Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition*, *20*(1), 125–151. https://doi.org/10.1016/0278-2626(92)90065-T

Demberg, V., & Sayeed, A. (2016). The frequency of rapid pupil dilations as a measure of linguistic processing difficulty. *PLOS ONE*, *11*(1), 1–29. https://doi.org/10.1371/journal.pone.0146194

Desai, M., Holder, J., Villarreal, C., Clark, N., Hoang, B., & Hamilton, L. S. (2021). Generalizable EEG encoding models with naturalistic audiovisual stimuli. *Journal of Neuroscience*, *41*(43), 8946–8962. https://doi.org/10.1523/JNEUROSCI.2891-20.2021

Dixon, M. L., Thiruchselvam, R., Todd, R., & Christoff, K. (2017). Emotion and the prefrontal cortex: An integrative review. *Psychological Bulletin*, *143*(10), 1033–1081. https://doi.org/10.1037/bul0000096

Dolcos, F., Iordan, A. D., & Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: A review of evidence from brain imaging investigations. *Journal of Cognitive Psychology*, *23*(6), 669–694. https://doi.org/10.1080/20445911.2011.594433

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*(1), 53–78. https://doi.org/10.1348/000712601162103

Duncan, J., Roberge, A., Fortier-Gauthier, U., Fiset, D., Blais, C., & Brisson, B. (2021). Dual task interference on early perceptual processing. *Attention, Perception, and Psychophysics*, *83*, 1777–1795. https://doi.org/10.3758/s13414-020-02158-0

Ekman, P. (1992b). An argument for basic emotions. *Cognition and Emotion*, *6*(3-4), 169–200. https://doi.org/10.1080/02699939208411068

Engel, A. K., & Fries, P. (2010). Beta-band oscillations – Signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165. https://doi.org/10.1016/j.conb.2010.02.015

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716. https://doi.org/10.1038/35094565

European Commission & Directorate General for Mobility and Transport. (2020). *Next steps towards 'Vision Zero' – EU road safety policy framework 2021-2030*. Publications Office of the European Union. https://doi.org/10.2832/391271

Fan, S., Dal Monte, O., & Chang, S. W. C. (2021). Levels of naturalism in social neuroscience research. *iScience*, *24*(7), 102702. https://doi.org/10.1016/j.isci.2021.102702

Ferencová, N., Višňovcová, Z., Bona Olexová, L., & Tonhajzerová, I. (2021). Eye pupil – A window into central autonomic regulation via emotional/cognitive processing. *Physiological Research*, *70*(4), 669–682. https://doi.org/10.33549/physiolres.934749

Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition and Emotion*, *16*(3), 355–379.

Fox, K. C., Andrews-Hanna, J. R., Mills, C., Dixon, M. L., Markovic, J., Thompson, E., & Christoff, K. (2018). Affective neuroscience of self-generated thought. *Annals of the New York Academy of Sciences*, *1426*(1), 25–51. https://doi.org/10.1111/nyas.13740

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., & Robbins, T. W. (2022). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*, *47*(1), 72–89. https://doi.org/10.1038/s41386-021-01132-0

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 671–678. https://doi.org/10.1098/rstb.2006.2003

Frühholz, S., & Grandjean, D. (2013). Multiple subregions in superior temporal cortex are differentially sensitive to vocal expressions: A quantitative meta-analysis. *Neuroscience & Biobehavioral Reviews*, *37*(1), 24–35. https://doi.org/10.1016/j.neubiorev.2012.11.002

Fusar-Poli, P., Placentino, A., Carletti, F., Allen, P., Landi, P., Abbamonte, M., Barale, F., Perez, J., McGuire, P., & Politi, P. (2009). Laterality effect on emotional faces processing: ALE meta-analysis of evidence. *Neuroscience Letters*, *452*(3), 262–267. https://doi.org/10.1016/j.neulet.2009.01.065

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., Benedetti, F., Abbamonte, M., Gasparotti, R., Barale, F., Pérez, J., McGuire, P., & Politi, P. (2009). Functional atlas of emotional faces processing: A voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of Psychiatry and Neuroscience*, *34*(6), 418–432. https://doi.org/10.1503/jpn.090014

García-Pacios, J., Del Río, D., Villalobos, D., Ruiz-Vargas, J. M., & Maestú, F. (2015). Emotional interference-based forgetting in short-term memory. Cognitive inhibition of pleasant but not unpleasant biologically relevant distractors. *Frontiers in Psychology*, *6*, 582. https://doi.org/10.3389/fpsyg.2015.00582

García-Pacios, J., Garcés, P., Del Río, D., & Maestú, F. (2015). Early detection and late cognitive control of emotional distraction by the prefrontal cortex. *Scientific Reports*, *5*, 10046. https://doi.org/10.1038/srep10046

Gendolla, G. H. E., & Silvestrini, N. (2011). Smiles make it easier and so do frowns: Masked affective stimuli influence mental effort. *Emotion*, *11*(2), 320–328. https://doi.org/10.1037/a0022593

Goldberg, H., Preminger, S., & Malach, R. (2014). The emotion–action link? Naturalistic emotional stimuli preferentially activate the human dorsal visual stream. *NeuroImage*, *84*, 254–264. https://doi.org/10.1016/j.neuroimage.2013.08.032

Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The neural bases of emotion regulation: Reappraisal and suppression of negative emotion. *Biological Psychiatry*, *63*(6), 577–586. https://doi.org/10.1016/j.biopsych.2007.05.031

Grahek, I., Musslick, S., & Shenhav, A. (2020). A computational perspective on the roles of affect in cognitive control. *International Journal of Psychophysiology*, *151*, 25–34. https://doi.org/10.1016/j.ijpsycho.2020.02.001

Grandjean, D. (2021). Brain networks of emotional prosody processing. *Emotion Review*, *13*(1), 34–43. https://doi.org/10.1177/1754073919898522

Gray, J. R., Braver, T. S., & Raichle, M. E. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proceedings of the National Academy of Sciences*, *99*(6), 4115–4120. https://doi.org/10.1073/pnas.062381899

Grecucci, A., Giorgetta, C., Bonini, N., & Sanfey, A. G. (2013). Reappraising social emotions: The role of inferior frontal gyrus, temporo-parietal junction and insula in interpersonal emotion regulation. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00523

Grillon, C., Quispe-Escudero, D., Mathur, A., & Ernst, M. (2015). Mental fatigue impairs emotion regulation. *Emotion*, *15*(3), 383. https://doi.org/10.1037/emo0000058

Grimshaw, G. M., & Carmel, D. (2014). An asymmetric inhibition model of hemispheric differences in emotional processing. *Frontiers in Psychology*, *5*, 489. https://doi.org/10.3389/fpsyg.2014.00489

Grissmann, S., Faller, J., Scharinger, C., Spüler, M., & Gerjets, P. (2017). Electroencephalography based analysis of working memory load and affective valence in an n-back task with emotional stimuli. *Frontiers in Human Neuroscience*, *11*, 616. https://doi.org/10.3389/fnhum.2017.00616

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*(3), 271–299. https://doi.org/10.1037/1089-2680.2.3.271

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*(3), 281–291. https://doi.org/10.1017/S0048577201393198

Gross, J. J. (2015a). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, *26*(1), 1–26. https://doi.org/10.1080/1047840X.2014.940781

Gross, J. J. (2015b). The extended process model of emotion regulation: Elaborations, applications, and future directions. *Psychological Inquiry*, *26*(1), 130–137. https://doi.org/10.1080/1047840X.2015.989751

Gross, J. J., & Jazaieri, H. (2014). Emotion, emotion regulation, and psychopathology: An affective science perspective. *Clinical Psychological Science*, *2*(4), 387–401. https://doi.org/10.1177/2167702614536164

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 3–24). The Guilford Press.

Gross, J. (2019). Magnetoencephalography in cognitive neuroscience: A primer. *Neuron*, *104*(2), 189–204. https://doi.org/10.1016/j.neuron.2019.07.001

Gyurak, A., Gross, J. J., & Etkin, A. (2011). Explicit and implicit emotion regulation: A dual-process framework. *Cognition and Emotion*, *25*(3), 400–412. https://doi.org/10.1080/02699931.2010.544160

Hämäläinen, M. S., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography – Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*(2), 413–497. https://doi.org/10.1103/RevModPhys.65.413

Hamilton, L. S., & Huth, A. G. (2020). The revolution will not be controlled: Natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience*, *35*(5), 573–582. https://doi.org/10.1080/23273798.2018.1499946

Hansen, P., Kringelbach, M., & Salmelin, R. (2010). *MEG: An introduction to methods*. Oxford University Press.

Harmon-Jones, E., Gable, P. A., & Peterson, C. K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biological Psychology*, *84*(3), 451–462. https://doi.org/10.1016/j.biopsycho.2009.08.010

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, *16*(2), 114–121. https://doi.org/10.1016/j.tics.2011.12.007

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, *87*, 96–110. https://doi.org/10.1016/j.neuroimage.2013.10.067

Häusler, C. O., & Hanke, M. (2021). A studyforrest extension, an annotation of spoken language in the German dubbed movie ”Forrest Gump” and its audio-description. *F1000Research*, *10*, 54. https://doi.org/10.12688/f1000research.27621.1

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Held, M., Rieger, J. W., & Borst, J. P. (2024). Multitasking while driving: Central bottleneck or problem state interference? *Human Factors*, *66*(5), 1564–1582. https://doi.org/10.1177/00187208221143857

Herlin, B., Navarro, V., & Dupont, S. (2021). The temporal pole: From anatomy to function — A literature appraisal. *Journal of Chemical Neuroanatomy*, *113*, 101925. https://doi.org/10.1016/j.jchemneu.2021.101925

Herrmann, C. S., Fründ, I., & Lenz, D. (2010). Human gamma-band activity: A review on cognitive and behavioral correlates and network models. *Neuroscience & Biobehavioral Reviews*, *34*(7), 981–992. https://doi.org/10.1016/j.neubiorev.2009.09.001

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. https://doi.org/10.1038/nrn2113

Higgins, E. T. (2006). Value from hedonic experience and engagement. *Psychological Review*, *113*(3), 439–460. https://doi.org/10.1037/0033-295X.113.3.439

Hogeveen, J., Bird, G., Chau, A., Krueger, F., & Grafman, J. (2016). Acquired alexithymia following damage to the anterior insula. *Neuropsychologia*, *82*, 142–148. https://doi.org/10.1016/j.neuropsychologia.2016.01.021

Holdgraf, C. R., Rieger, J. W., Micheli, C., Martin, S., Knight, R. T., & Theunissen, F. E. (2017). Encoding and decoding models in cognitive electrophysiology. *Frontiers in Systems Neuroscience*, *11*, 61. https://doi.org/10.3389/fnsys.2017.00061

Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020). The ‘real-world approach’ and its problems: A critique of the term ecological validity. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00721

Hooker, C. I., & Knight, R. T. (2006). The role of lateral orbitofrontal cortex in the inhibitory control of emotion. In *The orbitofrontal cortex*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198565741.003.0012

Huang, W., Wu, W., Lucas, M. V., Huang, H., Wen, Z., & Li, Y. (2023). Neurofeedback training with an electroencephalogram-based brain-computer interface enhances emotion regulation. *IEEE Transactions on Affective Computing*, *14*(2), 998–1010. https://doi.org/10.1109/TAFFC.2021.3134183

Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, *19*(3), 126–132. https://doi.org/10.1016/j.tics.2015.01.004

Jamil, M. H., Park, W., & Eid, M. (2021). Emotional responses to watching and touching 3D emotional face in a virtual environment. *Virtual Reality*, *25*(2), 553–564. https://doi.org/10.1007/s10055-020-00473-3

Jamil, N., Belkacem, A. N., Ouhbi, S., & Guger, C. (2021). Cognitive and affective brain–computer interfaces for improving learning strategies and enhancing student capabilities: A systematic literature review. *IEEE Access*, *9*, 134122–134147. https://doi.org/10.1109/ACCESS.2021.3115263

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186. https://doi.org/10.3389/fnhum.2010.00186

Jokisch, D., & Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *The Journal of Neuroscience*, *27*(12), 3244–3251. https://doi.org/10.1523/JNEUROSCI.5399-06.2007

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213–233. https://doi.org/10.1007/s11065-007-9040-z

Kalisch, R. (2009). The functional neuroanatomy of reappraisal: Time matters. *Neuroscience & Biobehavioral Reviews*, *33*(8), 1215–1226. https://doi.org/10.1016/j.neubiorev.2009.06.003

Kang, J.-H., Jeong, J. W., Kim, H. T., Kim, S. H., & Kim, S.-P. (2014). Representation of cognitive reappraisal goals in frontal gamma oscillations. *PLOS ONE*, *9*(11), 1–10. https://doi.org/10.1371/journal.pone.0113375

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109–2128. https://doi.org/10.1098/rstb.2006.1934

Karnath, H.-O. (2001). New insights into the functions of the superior temporal cortex. *Nature Reviews Neuroscience*, *2*(8), 568–576. https://doi.org/10.1038/35086057

Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLOS Biology*, *16*(3), 1–19. https://doi.org/10.1371/journal.pbio.2004473

Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective and Behavioral Neuroscience*, *6*(2), 110–126. https://doi.org/10.3758/CABN.6.2.110

Kim, H. (2019). Neural activity during working memory encoding, maintenance, and retrieval: A network-based model and meta-analysis. *Human Brain Mapping*, *40*(17), 4912–4933. https://doi.org/10.1002/hbm.24747

Kinner, V. L., Kuchinke, L., Dierolf, A. M., Merz, C. J., Otto, T., & Wolf, O. T. (2017). What our eyes tell us about feelings: Tracking pupillary responses during emotion regulation processes. *Psychophysiology*, *54*(4), 508–518. https://doi.org/10.1111/psyp.12816

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302*(5648), 1181–1185. https://doi.org/10.1126/science.1088545

Koole, S. L. (2010). The psychology of emotion regulation: An integrative review. In J. De Houwer & D. Hermans (Eds.), *Cognition and emotion: Reviews of current research and theories* (pp. 128–167). Psychology Press.

Krall, S. C., Rottschy, C., Oberwelland, E., Bzdok, D., Fox, P. T., Eickhoff, S. B., Fink, G. R., & Konrad, K. (2015). The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Structure and Function*, *220*(2), 587–604. https://doi.org/10.1007/s00429-014-0803-z

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure and Function*, *214*(5), 519–534. https://doi.org/10.1007/s00429-010-0255-z

Labs, A., Reich, T., Schulenburg, H., Boennen, M., Gehrke, M., Golz, M., Hartigs, B., Hoffmann, N., Keil, S., Perlow, M., Peukmann, A. K., Rabe, L. N., von Sobbe, F.-R., & Hanke, M. (2015). Portrayed emotions in the movie "Forrest Gump". *F1000Research*, *4*, 92. https://doi.org/10.12688/f1000research.6230.1

Lamm, C., & Singer, T. (2010). The role of anterior insular cortex in social emotions. *Brain Structure and Function*, *214*(5), 579–591. https://doi.org/10.1007/s00429-010-0251-3

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*(3), 143–148. https://doi.org/10.1177/0963721410370295

Lavie, N., & de Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review*, *12*(4), 669–674. https://doi.org/10.3758/BF03196756

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

LeDoux, J. (2012). Rethinking the emotional brain. *Neuron*, *73*(4), 653–676. https://doi.org/10.1016/j.neuron.2012.02.004

Leopold, D. A., Bondar, I. V., & Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature*, *442*(7102), 572–575. https://doi.org/10.1038/nature04951

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*(1), 89–94. https://doi.org/10.1038/82947

Lincoln, T. M., Schulze, L., & Renneberg, B. (2022). The role of emotion regulation in the characterization, development and treatment of psychopathology. *Nature Reviews Psychology*, *1*(5), 272–286. https://doi.org/10.1038/s44159-022-00040-4

Lindström, B. R., & Bohlin, G. (2011). Emotion processing facilitates working memory performance. *Cognition and Emotion*, *25*(7), 1196–1204. https://doi.org/10.1080/02699931.2010.527703

Lingelbach, K., Gado, S., Wirzberger, M., & Vukelić, M. (2023). Workload-dependent hemispheric asymmetries during the emotion-cognition interaction: A close-to-naturalistic fNIRS study. *Frontiers in Neuroergonomics*, *4*. https://doi.org/10.3389/fnrgo.2023.1273810

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157. https://doi.org/10.1038/35084005

Lopes da Silva, F. (2013). EEG and MEG: Relevance to neuroscience. *Neuron*, *80*(5), 1112–1128. https://doi.org/10.1016/j.neuron.2013.10.017

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marsicano, G., Bertini, C., & Ronconi, L. (2024). Decoding cognition in neurodevelopmental, psychiatric and neurological conditions with multivariate pattern analysis of EEG data. *Neuroscience & Biobehavioral Reviews*, *164*, 105795. https://doi.org/10.1016/j.neubiorev.2024.105795

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function*, *214*(5), 655–667. https://doi.org/10.1007/s00429-010-0262-0

Michels, L., Bucher, K., Lüchinger, R., Klaver, P., Martin, E., Jeanmonod, D., & Brandeis, D. (2010). Simultaneous EEG-fMRI during a working memory task: Modulations in low and high frequency bands. *PLoS ONE*, *5*(4), e10298. https://doi.org/10.1371/journal.pone.0010298

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1309–1316.

Monroe, J. F., Griffin, M., Pinkham, A., Loughead, J., Gur, R. C., Roberts, T. P. L., & Edgar, J. C. (2013). The fusiform response to faces: Explicit versus implicit processing of emotion. *Human Brain Mapping*, *34*(1), 1–11. https://doi.org/10.1002/hbm.21406

Mothes-Lasch, M., Miltner, W. H., & Straube, T. (2012). Processing of angry voices is modulated by visual load. *NeuroImage*, *63*(1), 485–490. https://doi.org/10.1016/j.neuroimage.2012.07.005

Murali, P. K., Kaboli, M., & Dahiya, R. (2022). Intelligent in-vehicle interaction technologies. *Advanced Intelligent Systems*, *4*(2), 2100122. https://doi.org/10.1002/aisy.202100122

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory – Where are we now, and where should we go next? *Psychonomic Bulletin & Review*, *23*, 1316–1340. https://doi.org/10.3758/s13423-015-0982-5

National Center for Statistics and Analysis. (2025, April). *Distracted driving in 2023* (Research report No. DOT HS 813 703; retrieved July 30, 2025). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813703

Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, *12*(2), 1211–1279. https://doi.org/10.3390/s120201211

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*(5827), 1002–1005. https://doi.org/10.1126/science.1136930

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1–18, Vol. 4). Springer US. https://doi.org/10.1007/978-1-4757-0629-1_1

Nummenmaa, L., Saarimäki, H., Glerean, E., Gotsopoulos, A., Jääskeläinen, I. P., Hari, R., & Sams, M. (2014). Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. *NeuroImage*, *102*, 498–509. https://doi.org/10.1016/j.neuroimage.2014.07.063

Oberauer, K. (2019). Working memory and attention – A conceptual analysis and review. *Journal of Cognition*, *2*(1), 36. https://doi.org/10.5334/joc.58

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–249. https://doi.org/10.1016/j.tics.2005.03.010

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: A synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, *1251*(1), 1–24. https://doi.org/10.1111/j.1749-6632.2012.06751.x

Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: A review of findings on social and emotional processing. *Brain*, *130*(7), 1718–1731. https://doi.org/10.1093/brain/awm052

Olsson, A., & Ochsner, K. N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences*, *12*(2), 65–71. https://doi.org/10.1016/j.tics.2007.11.010

Ou, J., & Law, S.-P. (2019). Top-down and bottom-up mechanisms as reflected by beta and gamma oscillations in speech perception: An individual-difference approach. *Brain and Language*, *199*, 104700. https://doi.org/10.1016/j.bandl.2019.104700

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148–158. https://doi.org/10.1038/nrn2317

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, *17*(4), 292–299. https://doi.org/10.1111/j.1467-9280.2006.01701.x

Pizzo, F., Roehri, N., Villalon, S. M., Trébuchon, A., Chen, S., Lagarde, S., Carron, R., Gavaret, M., Giusiano, B., McGonigal, A., Bartolomei, F., Badier, J. M., & Bénar, C. G. (2019). Deep brain activities can be detected with magnetoencephalography. *Nature Communications*, *10*(1), 971. https://doi.org/10.1038/s41467-019-08665-5

Popov, T., Steffen, A., Weisz, N., Miller, G. A., & Rockstroh, B. (2012). Cross-frequency dynamics of neuromagnetic oscillatory activity: Two mechanisms of emotion regulation. *Psychophysiology*, *49*(12), 1545–1557. https://doi.org/10.1111/j.1469-8986.2012.01484.x

Pruessner, L., Barnow, S., Holt, D. V., Joormann, J., & Schulze, K. (2020). A cognitive control framework for understanding emotion regulation flexibility. *Emotion*, *20*(1), 21–29. https://doi.org/10.1037/emo0000658

Quirk, G. J., & Beer, J. S. (2006). Prefrontal involvement in the regulation of emotion: Convergence of rat and human studies. *Current Opinion in Neurobiology*, *16*(6), 723–727. https://doi.org/10.1016/j.conb.2006.07.004

Redcay, E., & Moraczewski, D. (2020). Social cognition in context: A naturalistic imaging approach. *NeuroImage*, *216*, 116392. https://doi.org/https://doi.org/10.1016/j.neuroimage.2019.116392

Rellecke, J., Sommer, W., & Schacht, A. (2012). Does processing of emotional facial expressions depend on intention? Time-resolved evidence from event-related brain potentials. *Biological Psychology*, *90*(1), 23–32. https://doi.org/10.1016/j.biopsycho.2012.02.002

Rolls, E. T., Cheng, W., & Feng, J. (2020). The orbitofrontal cortex: Reward, emotion and depression. *Brain communications*, *2*(2), fcaa196. https://doi.org/10.1093/braincomms/fcaa196

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., Fox, P. T., & Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*, *60*(1), 830–846. https://doi.org/10.1016/j.neuroimage.2011.11.050

Ruiz, S., Buyukturkoglu, K., Rana, M., Birbaumer, N., & Sitaram, R. (2014). Real-time fMRI brain computer interfaces: Self-regulation of single brain regions to networks. *Biological Psychology*, *95*, 4–20. https://doi.org/10.1016/j.biopsycho.2013.04.010

Ruiz, S., Lee, S., Soekadar, S. R., Caria, A., Veit, R., Kircher, T., Birbaumer, N., & Sitaram, R. (2013). Acquired self-control of insula cortex modulates emotion recognition and brain network connectivity in schizophrenia. *Human Brain Mapping*, *34*(1), 200–212. https://doi.org/10.1002/hbm.21427

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805–819. https://doi.org/10.1037/0022-3514.76.5.805

Sabharwal, A., Szekely, A., Kotov, R., Mukherjee, P., Leung, H.-C., Barch, D. M., & Mohanty, A. (2016). Transdiagnostic neural markers of emotion-cognition interaction in psychotic disorders. *Journal of Abnormal Psychology*, *125*(7), 907–922. https://doi.org/10.1037/abn0000196

Said, C. P., Haxby, J. V., & Todorov, A. (2011). Brain systems for assessing the affective value of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1660–1670. https://doi.org/10.1098/rstb.2010.0351

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review*, *115*(1), 101–130. https://doi.org/10.1037/0033-295X.115.1.101

Savage, R. A., Lipp, O. V., Craig, B. M., Becker, S. I., & Horstmann, G. (2013). In search of the emotional face: Anger versus happiness superiority in visual search. *Emotion*, *13*(4), 758–768. https://doi.org/10.1037/a0031970

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological review*, *69*(5), 379. https://doi.org/10.1037/h0046234

Scherer, K. R., Schorr, A., & Johnstone, T. (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex*, *130*, 362–386. https://doi.org/10.1016/j.cortex.2020.06.010

Schirmer, A. (2018). Is the voice an auditory face? An ALE meta-analysis comparing vocal and facial emotion processing. *Social Cognitive and Affective Neuroscience*, *13*(1), 1–13. https://doi.org/10.1093/scan/nsx142

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009

Schlaffke, L., Lissek, S., Lenz, M., Juckel, G., Schultz, T., Tegenthoff, M., Schmidt-Wilcke, T., & Brüne, M. (2015). Shared and nonshared neural networks of cognitive and affective theory-of-mind: A neuroimaging study using cartoon picture stories. *Human Brain Mapping*, *36*(1), 29–39. https://doi.org/10.1002/hbm.22610

Schupp, H. T., Markus, J., Weike, A. I., & Hamm, A. O. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science*, *14*(1), 7–13. https://doi.org/10.1111/1467-9280.01411

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2021). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, *147*(3), 293. https://doi.org/10.1037/bul0000303

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition and Emotion*, *14*(4), 433–440. https://doi.org/10.1080/026999300402745

Schweizer, S., Satpute, A. B., Atzil, S., Field, A. P., Hitchcock, C., Black, M., Barrett, L. F., & Dalgleish, T. (2019). The impact of affective information on working memory: A pair of meta-analytic reviews of behavioral and neuroimaging evidence. *Psychological Bulletin*, *145*(6), 566–609. https://doi.org/10.1037/bul0000193

Scott, S. K. (2019). From speech and talkers to the social world: The neural processing of human spoken language. *Science*, *366*(6461), 58–62. https://doi.org/10.1126/science.aax0288

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, *27*(9), 2349–2356. https://doi.org/10.1523/JNEUROSCI.5587-06.2007

Shepherd, J. (2023). Conscious cognitive effort in cognitive control. *Wiley Interdisciplinary Reviews: Cognitive Science*, *14*(2), e1629.

Sheppes, G. (2020). Transcending the “good & bad” and “here & now” in emotion regulation: Costs and benefits of strategies across regulatory stages. In *Advances in experimental social psychology* (pp. 185–236, Vol. 61). Elsevier. https://doi.org/10.1016/bs.aesp.2019.09.003

Sheppes, G., Suri, G., & Gross, J. J. (2015). Emotion regulation and psychopathology. *Annual Review of Clinical Psychology*, *11*, 379–405. https://doi.org/10.1146/annurev-clinpsy-032814-112739

Skaramagkas, V., Giannakakis, G., Ktistakis, E., Manousos, D., Karatzanis, I., Tachos, N. S., Tripoliti, E., Marias, K., Fotiadis, D. I., & Tsiknakis, M. (2021). Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Reviews in Biomedical Engineering*, *16*, 260–277. https://doi.org/10.1109/RBME.2021.3066072

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends in Cognitive Sciences*, *23*(8), 699–714. https://doi.org/10.1016/j.tics.2019.05.004

Sonkusare, S., Qiong, D., Zhao, Y., Liu, W., Chen, H., Zhang, L., Li, M., Yang, K., & Wang, R. (2023). Frequency dependent emotion differentiation and directional coupling in amygdala, orbitofrontal and medial prefrontal cortex network with intracranial recordings. *Molecular Psychiatry*, *28*(4), 1636–1646. https://doi.org/10.1038/s41380-022-01883-2

Spiering, L., & Dimigen, O. (2024). (Micro)saccade-related potentials during face recognition: A study combining EEG, eye-tracking, and deconvolution modeling. *Attention, Perception, & Psychophysics*, 1–22. https://doi.org/10.3758/s13414-024-02846-1

Sweller, J., van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296. https://doi.org/10.1023/A:1022193728205

Tanaka, M., He, Z., Han, S., & Battaglia, S. (2025). Editorial: Noninvasive brain stimulation: A promising approach to study and improve emotion regulation. *Frontiers in Behavioral Neuroscience*, *19*. https://doi.org/10.3389/fnbeh.2025.1633936

Tavares, T. P., Logie, K., & Mitchell, D. G. (2016). Opposing effects of perceptual versus working memory load on emotional distraction. *Experimental Brain Research*, *234*, 2945–2956. https://doi.org/10.1007/s00221-016-4697-2

Tian, F., Hua, M., Zhang, W., Li, Y., & Yang, X. (2021). Emotional arousal in 2D versus 3D virtual reality environments. *PLOS ONE*, *16*(9), 1–17. https://doi.org/10.1371/journal.pone.0256211

Torre, J. B., & Lieberman, M. D. (2018). Putting feelings into words: Affect labeling as implicit emotion regulation. *Emotion Review*, *10*(2), 116–124. https://doi.org/10.1177/1754073917742706

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review*, *25*(6), 2005–2015. https://doi.org/10.3758/s13423-018-1432-y

van Dillen, L. F., & Derks, B. (2012). Working memory load reduces facilitated processing of threatening faces: An ERP study. *Emotion*, *12*(6), 1340–1349. https://doi.org/10.1037/a0028624

van Dillen, L. F., & Koole, S. L. (2009). How automatic is “automatic vigilance”? The role of working memory in attentional interference of negative information. *Cognition and Emotion*, *23*(6), 1106–1117. https://doi.org/10.1080/02699930802338178

van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*(3), 829–858. https://doi.org/10.1002/hbm.20547

van Pelt, S., Heil, L., Kwisthout, J., Ondobaka, S., van Rooij, I., & Bekkering, H. (2016). Beta- and gamma-band activity reflect predictive coding in the processing of causal events. *Social Cognitive and Affective Neuroscience*, *11*(6), 973–980. https://doi.org/10.1093/scan/nsw017

Vogels, J., Demberg, V., & Kray, J. (2018). The index of cognitive activity as a measure of cognitive processing load in dual task settings. *Frontiers in Psychology*, *9*, 2276. https://doi.org/10.3389/fpsyg.2018.02276

Vrba, J., & Robinson, S. E. (2001). Signal processing in magnetoencephalography. *Methods*, *25*(2), 249–271. https://doi.org/10.1006/meth.2001.1238

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*, *45*(1), 174–194. https://doi.org/10.1016/j.neuropsychologia.2006.06.003

Wagner, D. D., & Heatherton, T. F. (2012). Self-regulatory depletion increases emotional reactivity in the amygdala. *Social Cognitive and Affective Neuroscience*, *8*(4), 410–417. https://doi.org/10.1093/scan/nss082

Wang, J., Yang, Z., Klugah-Brown, B., Zhang, T., Yang, J., Yuan, J., & Biswal, B. B. (2024). The critical mediating roles of the middle temporal gyrus and ventrolateral prefrontal cortex in the dynamic processing of interpersonal emotion regulation. *NeuroImage*, *300*, 120789. https://doi.org/10.1016/j.neuroimage.2024.120789

Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, *90*(3), 1195–1268. https://doi.org/10.1152/physrev.00035.2008

Wang, Z., Li, S., & Shen, G. (2025). Advanced sensory hardware for intelligent eye-machine interfacing: From wearables to bionics. *Advanced Functional Materials*, 2503519. https://doi.org/https://doi.org/10.1002/adfm.202503519

Weidner, E. M., Moratti, S., Schindler, S., Grewe, P., Bien, C. G., & Kissler, J. (2024). Amygdala and cortical gamma-band responses to emotional faces are modulated by attention to valence. *Psychophysiology*, e14512. https://doi.org/10.1111/psyp.14512

Weidner, E. M., Schindler, S., Grewe, P., Moratti, S., Bien, C. G., & Kissler, J. (2022). Emotion and attention in face processing: Complementary evidence from surface event-related potentials and intracranial amygdala recordings. *Biological Psychology*, *173*, 108399. https://doi.org/10.1016/j.biopsycho.2022.108399

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Wickens, C. D. (2021). Attention: Theory, principles, models and applications. *International Journal of Human–Computer Interaction*, *37*(5), 403–417. https://doi.org/10.1080/10447318.2021.1874741

Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and Hearing*, *36*(4), 153–165. https://doi.org/10.1097/AUD.0000000000000145

Winston, J. S., O'Doherty, J., & Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage*, *20*(1), 84–97. https://doi.org/10.1016/s1053-8119(03)00303-3

Xu, Q., Ye, C., Gu, S., Hu, Z., Lei, Y., Li, X., Huang, L., & Liu, Q. (2021). Negative and positive bias for emotional faces: Evidence from the attention and working memory paradigms. *Neural Plasticity*, 8851066. https://doi.org/10.1155/2021/8851066

Yang, E., Milisav, F., Kopal, J., Holmes, A. J., Mitsis, G. D., Misic, B., Finn, E. S., & Bzdok, D. (2023). The default network dominates neural responses to evolving movie stories. *Nature Communications*, *14*(1), 4197. https://doi.org/10.1038/s41467-023-39862-y

Yao, S., Becker, B., Geng, Y., Zhao, Z., Xu, X., Zhao, W., Ren, P., & Kendrick, K. M. (2016). Voluntary control of anterior insula and its functional connections is feedback-independent and increases pain empathy. *NeuroImage*, *130*, 230–240. https://doi.org/10.1016/j.neuroimage.2016.02.035

Yin, Y.-Y., Yan, J.-Z., Lai, S.-X., Wei, Q.-Q., Sun, S.-R., Zhang, L.-M., & Li, Y.-F. (2024). Gamma oscillations in the mPFC: A potential predictive biomarker of depression and antidepressant effects. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *129*, 110893. https://doi.org/10.1016/j.pnpbp.2023.110893

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics*, *58*(1), 1–17. https://doi.org/10.1080/00140139.2014.956151

Zander, T. O., & Kothe, C. (2011). Towards passive brain–computer interfaces: Applying brain–computer interface technology to human–machine systems in general. *Journal of Neural Engineering*, *8*(2), 025005. https://doi.org/10.1088/1741-2560/8/2/025005

Zeelenberg, R., Wagenmakers, E.-J., & Rotteveel, M. (2006). The impact of emotion on perception: Bias or enhanced processing? *Psychological Science*, *17*(4), 287–291. https://doi.org/10.1111/j.1467-9280.2006.01700.x

Zhang, J., Bürkner, P.-C., Kiesel, A., & Dignath, D. (2023). How emotional stimuli modulate cognitive control: A meta-analytic review of studies with conflict tasks. *Psychological Bulletin*, *149*(1-2), 25–66. https://doi.org/10.1037/bul0000389

Zhang, R., Deng, H., & Xiao, X. (2024). The insular cortex: An interface between sensation, emotion and cognition. *Neuroscience Bulletin*, *40*(11), 1763–1773. https://doi.org/10.1007/s12264-024-01211-4

Zioga, I., Weissbart, H., Lewis, A. G., Haegens, S., & Martin, A. E. (2023). Naturalistic spoken language comprehension is supported by alpha and beta oscillations. *Journal of Neuroscience*, *43*(20), 3718–3732. https://doi.org/10.1523/JNEUROSCI.1500-22.2023

6 Supplementary Material for Study I

Neurophysiological Basis of Emotional Face Perception and Working Memory Load in a Dual-Task MEG Study

Katharina Lingelbach[1,2] and Jochem W. Rieger[1]

[1] Department of Psychology, University of Oldenburg, Oldenburg, Germany

[2] Fraunhofer Institute for Industrial Engineering IAO, Nobelstraße 12, 70569 Stuttgart, Germany

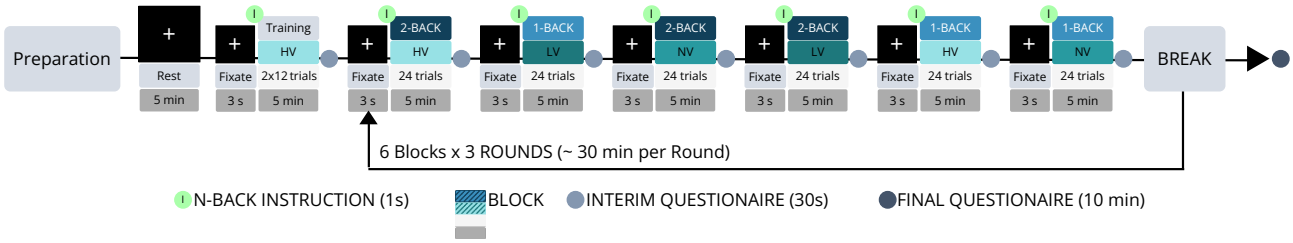
Supplementary Material to the article published as:

Lingelbach, K., & Rieger, J. W. (2025). Neurophysiological basis of emotional face perception and working memory load in a dual-task MEG study. *Human Brain Mapping, 46*(8), 1–25. https://doi.org/10.1002/hbm.70242

6.1 Supplementary Figure of the Experimental Procedure

Supplementary Figure 6.1

Overview of the Experiment Procedure



Note. The order of the conditions is an illustrative example. The provided times are approximated. Conditions were presented randomly within one round without repetition. LV: low valence; NV: neutral valence; HV: high valence; I: instruction.

6.2 Summary of the Behavioural Linear Mixed-Effects Models

Supplementary Table 6.1

Summary of the Behavioural Linear Mixed-Effects Models Analysing the Main Effects Emotional Expression and Working Memory (WM) Load, as Well as Their Interaction During the Facial Emotion Discrimination and N-Back Encoding/Retrieval Phase

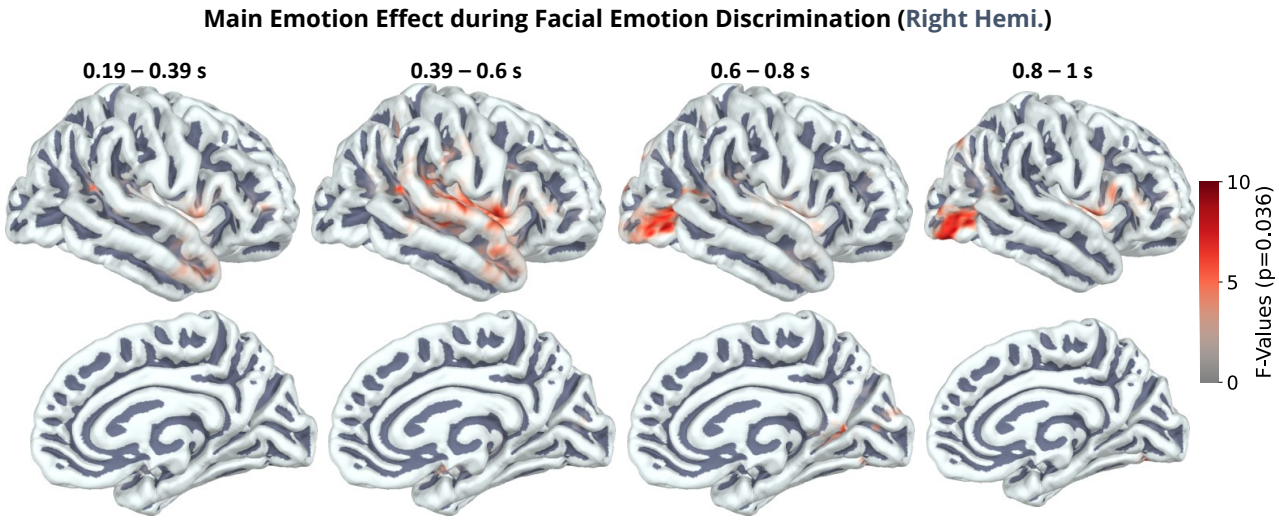
Measure	Effect	*df1*	*df2*	*F*	*p*
Performance					
Errors Facial Emotion Discrimination	Workload	230	1	0.05	.826
	Emotion	230	2	24.52	<.001***
	Workload * Emotion	230	2	0.30	.741
Errors N-Back Retrieval	Workload	230	1	68.27	<.001***
	Emotion	230	2	0.50	.607
	Workload * Emotion	230	2	0.01	.988
Reaction Time Facial Emotion Discrimination	Workload	230	1	9.91	.002**
	Emotion	230	2	15.27	<.001***
	Workload * Emotion	230	2	0.11	.898
Reaction Time N-Back Retrieval	Workload	230	1	15.21	<.001***
	Emotion	230	2	2.15	.118
	Workload * Emotion	230	2	0.28	.754
Subjective Ratings					
Valence (SAM Subscale)	Workload	230	1	6.27	.013*
	Emotion	230	2	18.80	<.001***
	Workload * Emotion	230	2	0.99	.373
Arousal (SAM Subscale)	Workload	230	1	41.08	<.001***
	Emotion	230	2	3.33	.038*
	Workload * Emotion	230	2	1.75	.175
Effort (NASA TLX Subscale)	Workload	230	1	75.71	<.001***
	Emotion	230	2	1.93	.148
	Workload * Emotion	230	2	0.93	.396

Note. Significance level: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

6.3 Spatiotemporal *F*-Statistic of the Permutation-Based Clustering

Supplementary Figure 6.2

Spatiotemporal F-statistic Cluster of the Main Effect Emotional Expression and the Temporally Evolving Spatial Distribution of Modulated Event-Related Magnetic Fields (ERFs)



Note. ERFs: event-related magnetic fields. Hemi.: hemisphere.

6.4 Exploratory Correlation Analyses Within and Across Spatiotemporal Event-Related Magnetic Fields Clusters

The spatiotemporal clustering of event-related magnetic fields (ERFs) does not permit inferring potential relationships between cluster regions and time intervals on a trial-by-trial level because it is based on averaged MEG activations across trials. This follow-up analysis focuses on relations within and between regions and time intervals involved in the process of facial emotion discrimination, as revealed by the ERF sub-clusters. The objective was to identify the processing stream across time intervals and ROIs for each condition.

For this purpose, we refined the spatiotemporal ERF clusters by taking only *t*-values corresponding to a significance threshold of $\alpha < .01$ ($t > 2.715$ and $t < -2.715$). Each cluster was segmented into four time intervals, and significant *t*-values within each time interval were averaged over time. Next, we extracted two spatially separated regions of interest (ROIs) maximally distinguishing the conditions of the comparison, as determined by their *t*-value. To obtain the primary peak (ROI 1), we drew a 10 mm radius around the maximum. To obtain the secondary peak (ROI 2), we applied the constraint that it must be located at least 20 mm from the primary peak. Thus, the secondary peak is the maximum when excluding all points in a 20 mm radius

around and within the primary peak. The first rule ensures that the selected ROI captures a sufficient number of vertices while still being focused on a functionally meaningful region. The second 20 mm separation rule between the primary and secondary peaks ensures that the two regions are spatially distinct. This is important to prevent artificial correlation that may arise due to point-spread function effects.

The ROIs per post-hoc comparison, cluster, and binned time interval from the significant ERF clusters (i.e., from each panel in Figure 2.5 in the main text) are depicted in the Supplementary Figure 6.3. MNI coordinates of the ROIs (i.e., peaks per bin), size of the defined ROI (in vertices), and associated anatomical labels based on the Automated Anatomical Labelling (AAL) atlas and Brodmann atlas (BA) are provided in the Supplementary Table 6.2.

Supplementary Table 6.2

Descriptive Parameters of the Regions of Interest (ROIs) for Each Comparison, Cluster per Hemisphere, and Time Interval of Interest (TOI)

Comparison	Cluster	Hemi.	TOI	Primary Peak				Secondary Peak			
				MNI Coordinates	Vertices in ROI	Label	BA	MNI Coordinates	Vertices in ROI	Label	BA
LV – NV	1	left	1	[-26, -92, 8]	38	OFA L	18	[-26, -93, -15]	7	Lingual L	18
			2	[-29, -92, -8]	38	OFA L	18	[-58, -17, 1]	15	Mid Temporal L	22
			3	[-35, -82, -14]	33	FFA L	19	[-24, -89, 4]	7	OFA L	18
			4	[-39, -85, -9]	38	OFA L	19	[-13, -97, -12]	12	Lingual L	18
		right	1	[29, -92, -8]	33	OFA R	18	[19, -81, -12]	8	Lingual R	18
			2	[55, 4, 3]	41	FO R	44	[39, -24, 0]	11	STS R	41
			3	[52, 0, 5]	39	FO R	44	[51, 30, 4]	16	IFG R	45
			4	[54, 2, 3]	40	FO R	44	[38, 0, 2]	10	Insula R	13
LV – HV	1	left	1	[-23, -98, 0]	36	OFA L	18	[-42, -80, -10]	8	OFA L	19
			2	[-38, -79, -13]	40	FFA L	19	[-29, -89, 13]	13	OFA L	19
			3	[-39, -85, -9]	38	OFA L	19	[-50, -20, 6]	10	STS L	41
			4	[-41, -85, 0]	41	OFA L	18	[-27, -60, 43]	22	Sup Parietal L	7
		right	1	[23, -94, -11]	34	OFA R	18	[57, -1, -26]	10	Mid Temporal R	21
			2	[57, -1, -26]	32	Mid Temporal R	21	[33, -88, -1]	7	OFA R	18
			3	[44, -76, -9]	37	OFA R	19	[53, -56, 10]	17	Mid Temporal R	37
			4	[41, -78, -13]	37	OFA R	19	[52, 3, 3]	11	FO R	44
	2	left	1	[-22, 30, 32]	45	Sup Frontal L	8	[-6, 3, 65]	10	SMA L	6
			2	[-41, 35, 22]	54	IFG L	46	[-14, 14, 57]	9	Sup Frontal L	6
			3	[-11, 28, 50]	30	Sup Frontal L	8	[-36, 47, 18]	10	Mid Frontal L	10
			4	[-15, 8, 60]	35	Sup Frontal L	6	[-29, 38, 19]	7	Mid Frontal L	10

Note. Labels were defined with the aid of the Automated Anatomical Labelling (AAL) atlas and Brodmann atlas. MNI: Montreal Neurological Institute space; BA: Brodmann area; L: left; R: right; LV: low valence; NV: neutral valence; HV: high valence; OFA: occipital face area; FFA: fusiform face area; Mid: middle; Sup: superior; FO: frontal operculum; SMA: supplementary motor area. LV – NV Clusters and LV – HV Cluster 1 with TOIs ranging from (1) 190 – 389 ms, (2) 390 – 600 ms, (3) 600 – 800 ms, (4) 800 – 1,000 ms. LV – HV Cluster 2 with TOIs ranging from (1) 400 – 550 ms, (2) 550 – 700 ms, (3) 700 – 850 ms and (4) 850 to 1,000 ms.

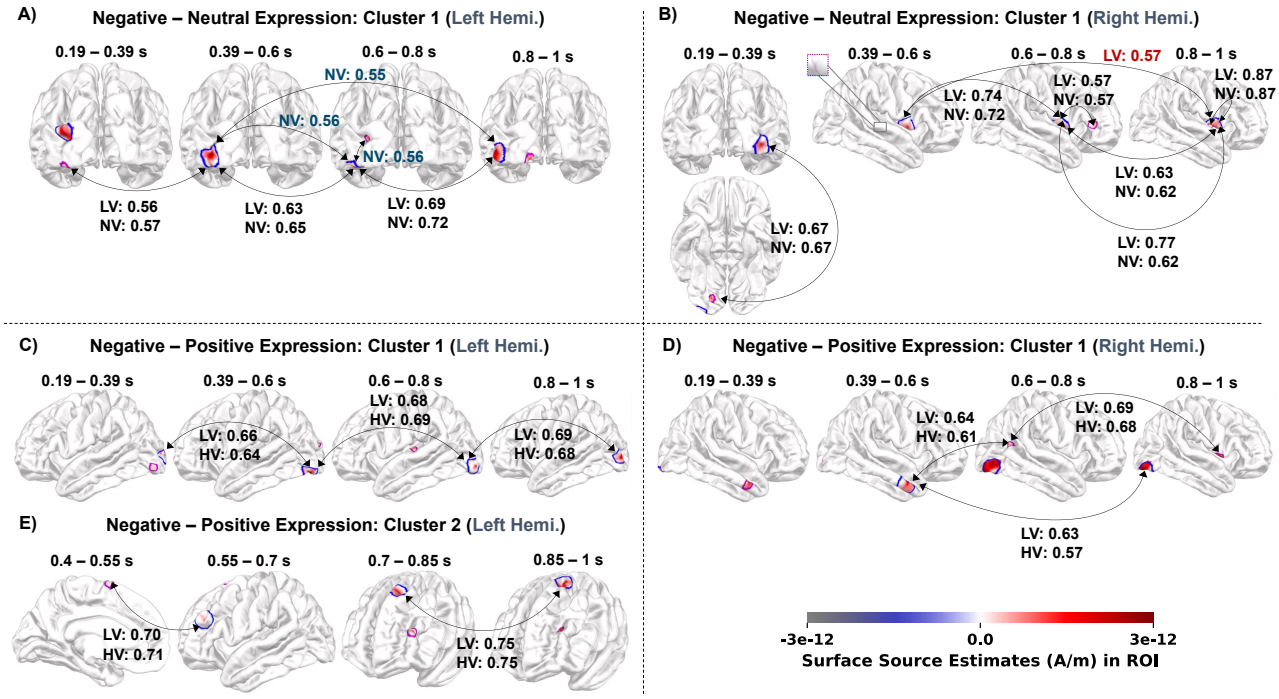
The source-space transformation was performed using the same approach as for the time-locked ERF and oscillatory power analyses. The Minimum Norm Estimate (MNE) was applied subject-wise to epoched data as inverse method, using an SNR value of 1 (Gramfort et al., 2013; Lin et al., 2006). We extracted the

individual source activity, which had been previously transformed into a common space, from the specified ROIs for each cluster, time interval, and condition. Source estimates per cluster, time interval, and ROI were correlated subject-wise (as compared to the group-level approach in the cluster analysis) using Spearman rank correlations (r_s). This allowed quantifying the spatiotemporal relationships within and between time intervals and clusters across trials in each subject.

To discern significant relationships between time intervals and clusters across conditions, we calculated the *t*-value of each correlation coefficient and computed the grand average of the *t*-values across participants per condition and correlation coefficient. Correlations with a mean *t*-value above the α-threshold corresponding to a Bonferroni-corrected p-value of 0.05 were defined as statistically significant relationships. The significance threshold was $-3.91 < t > 3.91$, for the correlation coefficients of the clusters LV – NV (LV: low valence; NV: neutral valence), and $-4.19 < t > 4.19$, for those of clusters LV – HV (HV: high valence).

Supplementary Figure 6.3

Regions of Interest (ROIs) Comprising the Primary Peak (Blue Outline) and Secondary Peak (Magenta Outline) per Comparison, Cluster, and Time Interval (see Figure 2.5)

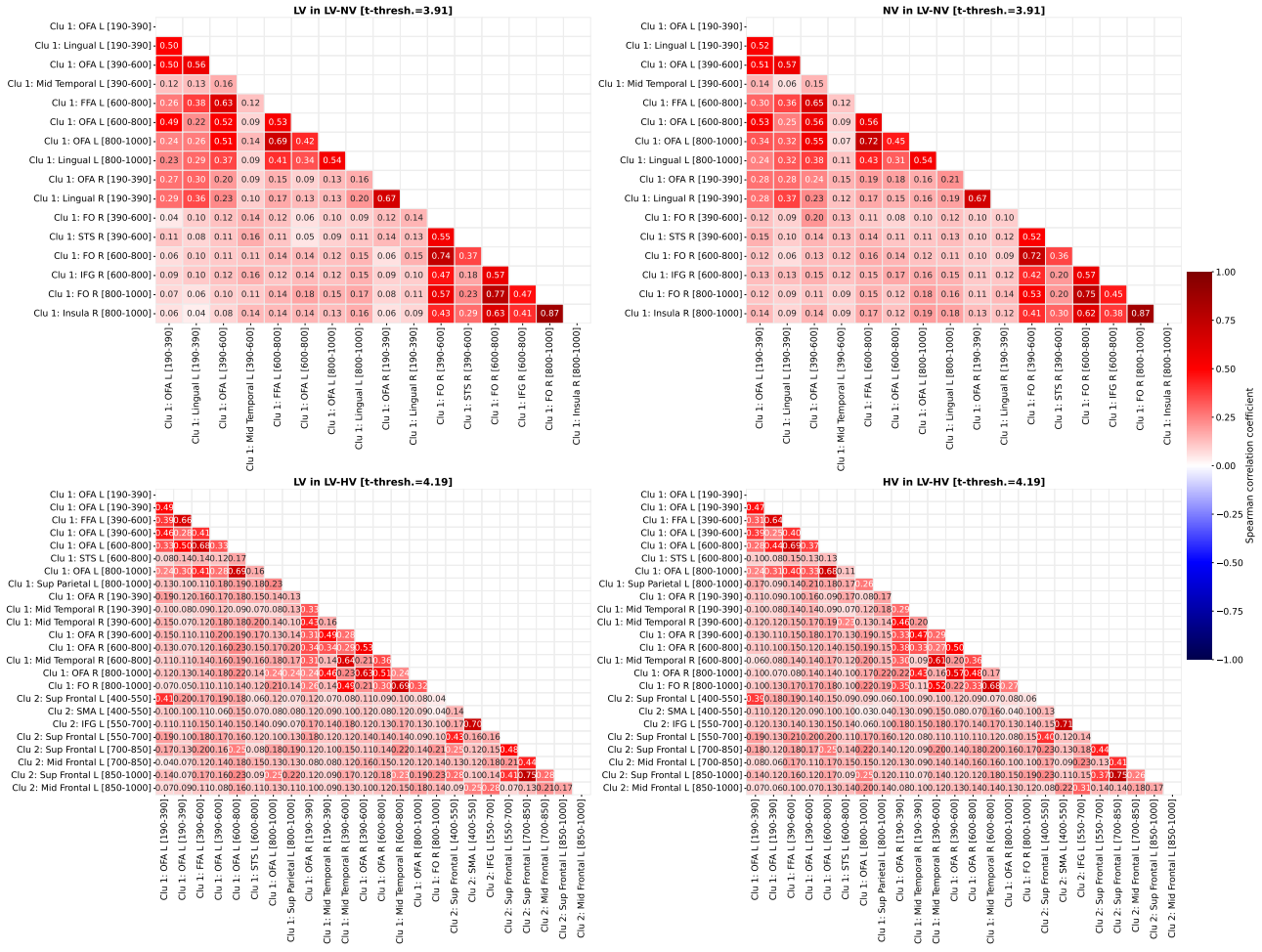


Note. (A) Caudal perspective for the left hemisphere and (B) caudal and inferior (180 – 390 ms), as well as lateral perspective (390 – 1000 ms) for the right hemisphere of the contrast negative – neutral emotional expressions (Figure 2.5A, B). Lateral perspective for the (C) left and (D) right hemisphere, as well as (E) medial (400 – 550 ms), lateral (550 – 700 ms), and frontal (700 – 1000 ms) perspective for the second left hemispheric cluster of the contrast negative – positive emotional expressions (Figure 2.5C – E). Significant spatiotemporal correlation coefficients within and among the ROIs and time intervals are labelled.

Our results revealed that significant correlations occurred mainly between temporally and spatially close ROIs (Supplementary Figure 6.3). All correlation matrices are provided in the Supplementary Figure 6.4. For all conditions, we observed strong correlations between ROIs located in face-specific occipital regions, that are the OFA and FFA, as well as lingual regions within the same and across adjacent time intervals.

Supplementary Figure 6.4

Correlation Matrices of the Trial-by-Trial Spearman Rank Correlations Between Regions of Interest (ROIs)



Note. ROIs were extracted from significant spatiotemporal event-related magnetic fields (ERFs) for each comparison, cluster per hemisphere, and time interval of interest (TOI; Figure 2.5). Labels were defined using the Automated Anatomical Labelling (AAL) atlas and Brodmann atlas. Clu: Cluster; L: left; R: right; LV: low valence; NV: neutral valence; HV: high valence; OFA: occipital face area; FFA: fusiform face area; Mid: middle; Sup: superior; FO: frontal operculum; SMA: supplementary motor area. LV – NV clusters and LV – HV Cluster 1 with TOIs ranging from (1) 190 – 389 ms, (2) 390 – 600 ms, (3) 600 – 800 ms, and (4) 800 – 1,000 ms. LV – HV Cluster 2 with TOIs ranging from (1) 400 – 550 ms, (2) 550 – 700 ms, (3) 700 – 850 ms, and (4) 850 – 1,000 ms.

Using the ROIs extracted from the negative – neutral expression contrast (Figure 2.5A, B), neutral

faces elicited additional correlation between OFA activation in early time intervals (390 – 600 ms) to later OFA activation (600 – 800 ms, $r_s = .56$; and 800 – 1000 ms, $r_s = .55$), and FFA activation (600 – 800 ms, $r_s = .56$; Supplementary Figure 6.3A). Furthermore, we observed a strong correlation of activation in the right FO across time intervals (390 – 600 ms to 600 – 800 ms, NV: $r_s = .72$, LV: $r_s = .74$), as well as co-activation of the right FO (600 – 800 ms) and right IFG within the same time interval (600 – 800 ms, NV and LV: $r_s = .57$; Supplementary Figure 6.3B). Moreover, the right FO was correlated with insular activation across time intervals (600 – 800 ms to 800 – 1000 ms, NV: $r_s = .62$, LV: $r_s = .63$), and within the final time interval (800 – 1000 ms, NV and LV: $r_s = .87$). For negative faces, we observed an additional correlation between early FO engagement (390 – 600 ms) and later insula activation (800 – 1000 ms) in the right hemisphere ($r_s = .57$; Supplementary Figure 6.3B).

Using the ROIs of the negative – positive expression contrast (Figure 2.5C – E), activation of the right mid-temporal sulcus (600 – 800 ms) was correlated with the right FO (800 – 1000 ms, HV: $r_s = .68$, LV: $r_s = .69$; Supplementary Figure 6.3D). Moreover, there were significant correlations in the second negative – positive expression cluster between the left supplementary motor area (SMA; 400 – 550 ms) and left IFG (550 – 700 ms, HV: $r_s = .71$, LV: $r_s = .70$), as well as within the superior frontal lobe in neighbouring time intervals (700 – 850 to 850 – 1000 ms, HV: $r_s = .75$, LV: $r_s = .75$; Supplementary Figure 6.3E).

The correlation analysis demonstrated that the processing stream for faces has few temporally and spatially distant connections within the spatiotemporal ERF clusters. Most connections were observed in face-processing-related occipital regions, the superior temporal sulcus, the insula, and inferior frontal regions at the same or adjacent time intervals. Since a correlational approach with few assumptions and no definition of direction was chosen, the possibility of moderators and other sources of activity cannot be excluded.

6.5 Exploratory Correlation Analyses of Evoked Oscillatory Power and Memory Performance

We investigated the role of working memory (WM) load-driven modulations in the alpha and beta band cluster and their relationship with high WM load performance (Bonnefond & Jensen, 2012; Proskovec, Wiesman, et al., 2019). This was done by correlating individual cluster-based alpha and low beta band power modulations (directional change HW – LW) with performance measures (reaction time and errors) in the high WM load n-back scenario. Periodic power represents putative oscillations associated with physiological, cognitive, and behavioural states (Donoghue et al., 2020). To extract periodic power, the aperiodic ($1/f$ like) component was removed from the power spectral density and averaged across epochs, for each participant, condition, and vertex. First, the $1/f$ decay was estimated without the knee parameter using the fooof algorithm (v. 1.1.0; Donoghue et al., 2020). Then, the $1/f$ component was subtracted from the original power spectral density (both in linear space without log-transformation; Gyurkovics et al., 2021b). The resulting periodic power was then averaged across band frequencies and vertices within the significant cluster. Spearman rank correlations (r_s) were computed, with *z*-standardised variables.

Supplementary Table 6.3

Correlation Analyses Between Directional Modulation (HW – LW) in Individual Periodic Alpha and Low Beta Cluster Power and N-Back Performance During High WM Load (Reaction Time and Errors)

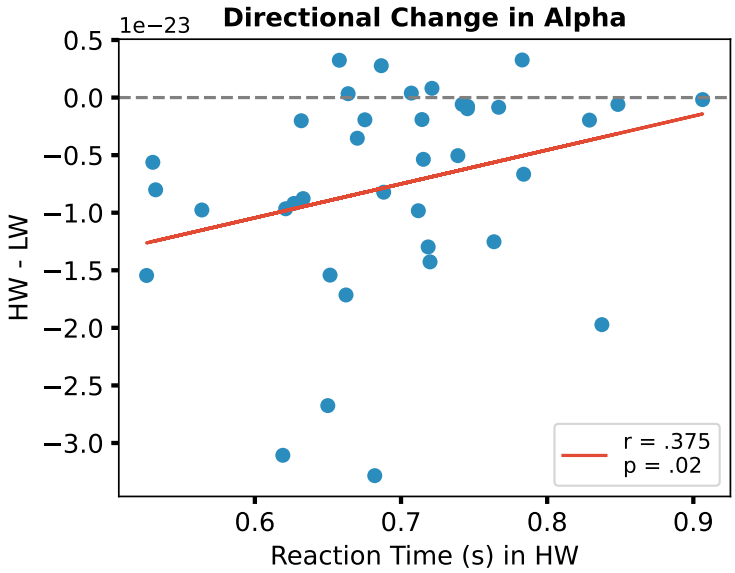
	Reaction Time in HW		Errors in HW	
Frequency Band Modulation	$r_{spearman}$	***p***	$r_{spearman}$	***p***
Alpha HW – LW	.375	.020*	-.032	.850
Low Beta HW – LW	.118	.482	-.095	.570

Note. HW: High WM load; LW: Low WM load; HW – LW: directional change; $r_{spearman}$: Spearman rank correlation; Significance level: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

The analysis revealed a positive relationship between periodic power modulation in the alpha band cluster and retrieval reaction time during the high WM load n-back subtask (Supplementary Table 6.3 and Supplementary Figure 6.5). There was a significant correlation between directional oscillatory changes in the alpha band cluster (HW – LW) and retrieval reaction time ($r_s = .375, p = .02$). No relationship was observed between alpha band modulation and retrieval accuracy, nor between beta band power modulation and WM task performance.

Supplementary Figure 6.5

Correlation Between Individual Periodic Working Memory (WM) Load-Based Alpha Band Modulation (i.e., Directional Oscillatory Changes HW – LW) and N-Back Retrieval Reaction Time Under High WM Load



Note. The analysis focuses on individual periodic alpha band power, averaged across cluster vertices identified through a permutation-based cluster analysis. HW: High WM load; LW: Low WM load; r: Spearman rank correlation.

Our results suggest that temporo-occipital alpha suppression (i.e., a decrease in alpha band power within the cluster) under high WM load serves as a compensatory mechanism to maintain task performance by enhancing attention allocation and sensory processing. However, this alpha band modulation does not correlate with n-back retrieval accuracy under high WM load (Supplementary Table 6.3), indicating that it is not directly involved in maintaining the to-be-retrieved information in WM. Importantly, the relationship between oscillatory modulation and n-back retrieval time was restricted to the cluster regions (global alpha power HW – LW x HW reaction time: $r_s = .235, p = .155$).

To conclude, WM load-driven modulation in the alpha band cluster likely represents a compensatory mechanism to enhance attention allocation and information processing under high WM load.

6.6 Summary of the Gaze-Related Linear Mixed-Effects Models

Supplementary Table 6.4

Summary of the Gaze-Related Linear Mixed-Effects Models Analysing the Main Effects Emotional Expression and WM Load as Well as Their Interaction During the Facial Emotion Discrimination and N-Back Encoding/Retrieval Phase

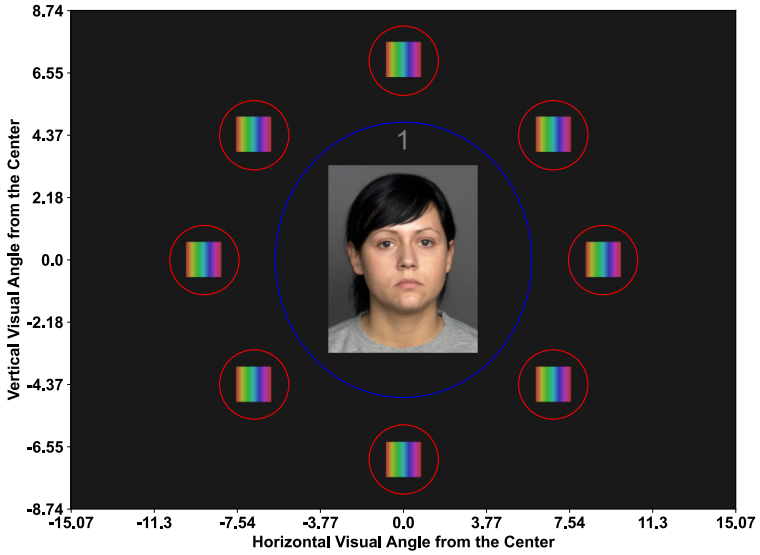
Facial Emotion Discrimination					
Mean Fixation Duration on Face	Workload	195	1	0.12	.731
	Emotion	195	2	4.10	.018*
	Workload * Emotion	195	2	1.19	.305
Count of Fixations on Face	Workload	195	1	2.87	.092
	Emotion	195	2	8.17	$<$.001***
	Workload * Emotion	195	2	0.90	.409
Mean Pupil Dilation in Fixations	Workload	195	1	78.99	$<$.001***
	Emotion	195	2	1.76	.176
	Workload * Emotion	195	2	4.46	.013*
Visuo-spatial Working Memory					
Mean Fixation Duration on Face	Workload	195	1	7.10	.008**
	Emotion	195	2	0.53	.590
	Workload * Emotion	195	2	0.11	.895
Count of Fixations on Face	Workload	195	1	0.43	.513
	Emotion	195	2	1.81	.167
	Workload * Emotion	195	2	0.87	.419
Mean Pupil Dilation in Fixations	Workload	195	1	16.65	$<$.001***
	Emotion	195	2	1.95	.145
	Workload * Emotion	195	2	1.81	.167
Onset of Fixation on Target Position Square	Workload	176.25	1	13.26	$<$.001***
	Emotion	177.21	2	1.65	.195
	Workload * Emotion	176.22	2	0.34	.714
Onset of Fixation on Target Colour Square	Workload	185.38	1	2.32	.129
	Emotion	184.24	2	0.21	.812
	Workload * Emotion	184.26	2	1.03	.359
Count of Fixation on Target Position Square	Workload	195	1	0.08	.777
	Emotion	195	2	1.06	.347
	Workload * Emotion	195	2	4.35	.014*
Count of Fixation on Target Colour Square	Workload	195	1	0.06	.814
	Emotion	195	2	0.69	.502
	Workload * Emotion	195	2	3.36	.037*

Note. Significance level: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

6.7 Effects of Dual-Task Phase on Gaze Behaviour

Supplementary Figure 6.6

Defined Regions of Interest (ROIs) Around the Face Stimuli (Blue Circle) and Squares (Red Circles) for the Fixation-Related Eye-Tracking Analysis



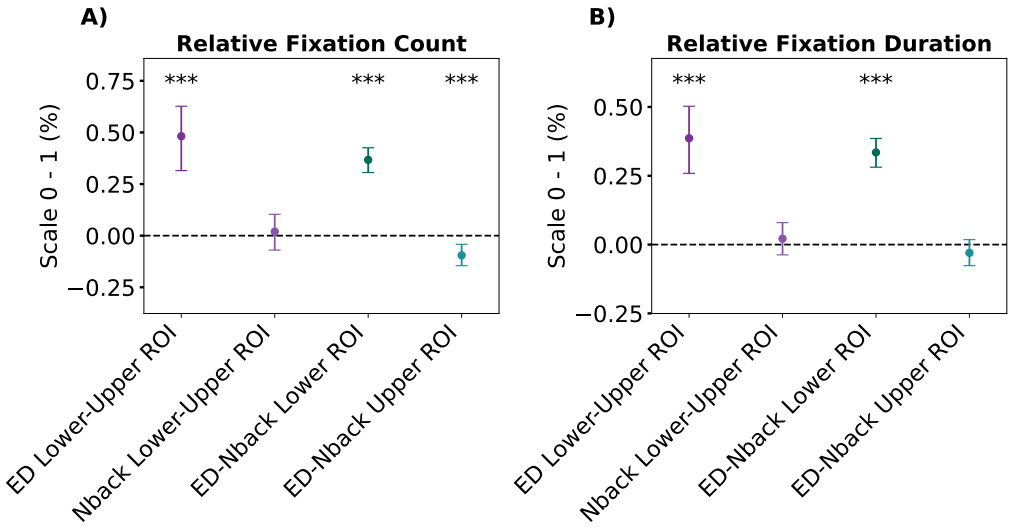
6.7.1 Fixations on Face Regions

To analyse whether one half of the face was preferred during fixations depending on the dual-task phase, we vertically divided the face ROI into an upper part, including the eyes, and a lower part, including the mouth. The relative fixation count and duration were extracted subject-wise for each emotional expression, facial region, and dual-task phase (i.e., during the emotion discrimination, ED; and n-back encoding/retrieval phase, N-Back). The three-way ANOVA with the factors emotional expression (HV, NV, LV) x facial region (upper, lower) x task phase (ED, N-Back) revealed a significant main effect of face region on both relative fixation count ($F_{(1,\ 468)} = 243.27, p < .001$) and duration ($F_{(1,\ 429)} = 252.76, p < .001$). There was no main effect of emotional expression, nor a significant interaction with emotional expression. Task phase had a significant main effect on relative fixation count ($F_{(1,\ 468)} = 71.03, p < .001$) and duration ($F_{(1,\ 429)} = 140.49, p < .001$).

Importantly, we found a significant interaction between face region and task phase, affecting both relative fixation count ($F_{(1,\ 468)} = 206.22$, $p < .001$) and duration ($F_{(1,\ 429)} = 202.66$, $p < .001$; Supplementary Figure 6.7). The interaction indicated that the preference for a facial region depended on the specific dual-task phase and subtask. The lower face region was favoured in fixations only during emotion discrimination (ED lower – upper ROI relative fixation count: $M = 0.482\ \%$; $95\ \%\ CI\ [0.315,\ 0.627]$; relative fixation duration: $M = 0.386\ \%$; $95\ \%\ CI\ [0.259,\ 0.503]$), but not during the n-back encoding/retrieval phase (N-Back lower – upper ROI relative fixation count: $M = 0.019\ \%$; $95\ \%\ CI\ [-0.070,\ 0.103]$; relative fixation duration: $M = 0.021\ \%$; $95\ \%\ CI\ [-0.037,\ 0.080]$; Supplementary Figure 6.7). When comparing task phase effects per region, more fixations were positioned in the lower face region during the emotion discrimination compared to the n-back encoding/retrieval phase (ED – N-Back lower ROI: $M = 0.367\ \%$; $95\ \%\ CI\ [0.306,\ 0.426]$), and significantly fewer fixations were located in the upper face region during emotion discrimination (ED – N-Back upper ROI: $M = -0.095\ \%$; $95\ \%\ CI\ [-0.145,\ -0.042]$). Similarly, fixation duration was significantly longer on the lower face region (ED – N-Back lower ROI: $M = 0.335\ \%$; $95\ \%\ CI\ [0.281,\ 0.386]$). No significant difference was found for the upper face region (ED – N-Back upper ROI: $M = -0.030\ \%$; $95\ \%\ CI\ [-0.077,\ 0.018]$; Supplementary Figure 6.7).

Supplementary Figure 6.7

Exploratory Interaction Comparisons of Task Phase (Emotion Discrimination, N-Back Encoding/Retrieval) and Face Region (Upper and Lower Half of the Face) for Relative Fixation Count and Duration



Note. Coloured dots and error bars represent the bootstrapped grand averages and their Bonferroni-corrected 2.5th and 97.5th confidence interval (CI) across participants. ED: emotion discrimination; Nback: n-back encoding/retrieval phase. Significance level from the linear mixed models: *** for $p < .001$, ** for $p < .01$, * for $p < .05$.

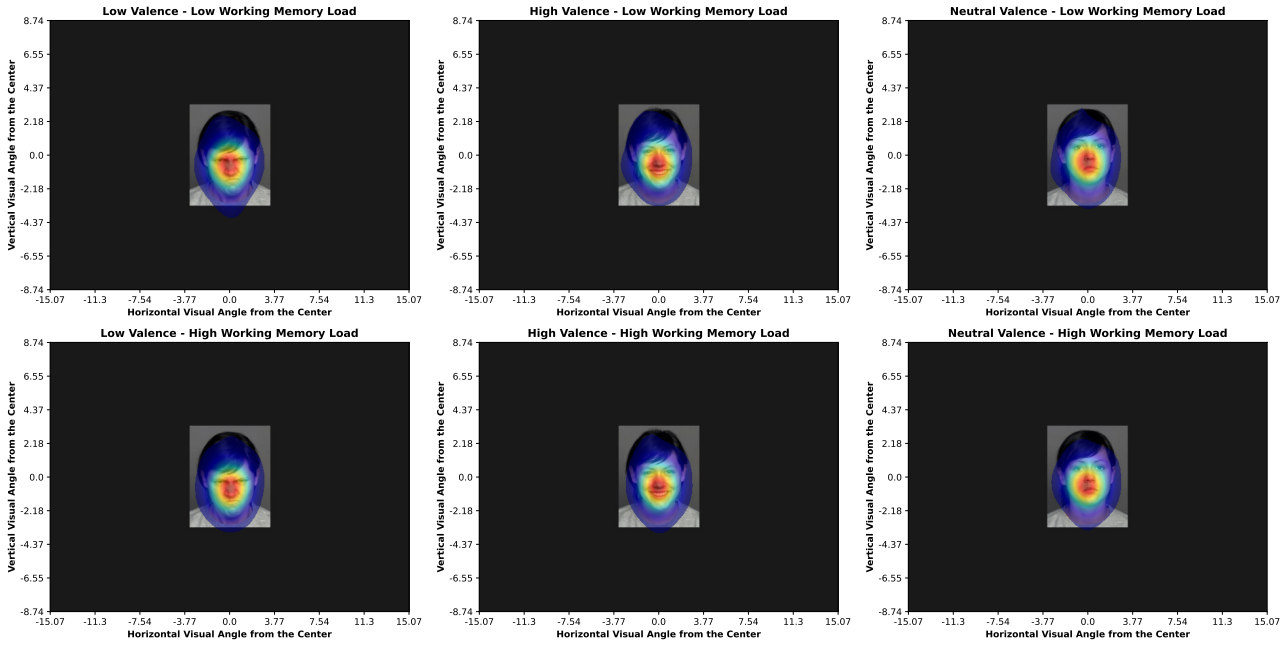
6.7.2 Face Fixations and Pupil Dilation

To compare the two dual-task phases, we extracted pupil dilation as well as the relative fixation count and duration on the face during the emotion discrimination (ED) and n-back encoding/retrieval phase (N-Back; full face ROI in Supplementary Figure 6.6; cf. Supplementary Analysis 6.7.1).

In comparing eye behaviour across the two dual-task phases, we found a significant main effect of task phase on relative fixation count ($F_{(1,\ 429)} = 1566.05$, $p < .001$) and duration ($F_{(1,\ 429)} = 1203.29$, $p < .001$) on the ROI of the face stimulus, as well as pupil dilation during fixations ($F_{(1,\ 429)} = 90.24$, $p < .001$). During the emotion discrimination, faces were fixated more frequently (ED – N-Back: $M = 30.460\ \%$; $95\ \%\ CI\ [26.485,\ 34.490]$) and for longer durations (ED – N-Back: $M = 27.188\ \%$; $95\ \%\ CI\ [23.928,\ 30.465]$), regardless of the displayed emotion or current WM load level. Pupil dilation was increased during emotion discrimination compared to the n-back encoding/retrieval phase (ED – N-Back: $M = 0.145$ mm; $95\ \%\ CI\ [0.113,\ 0.177]$).

Supplementary Figure 6.8

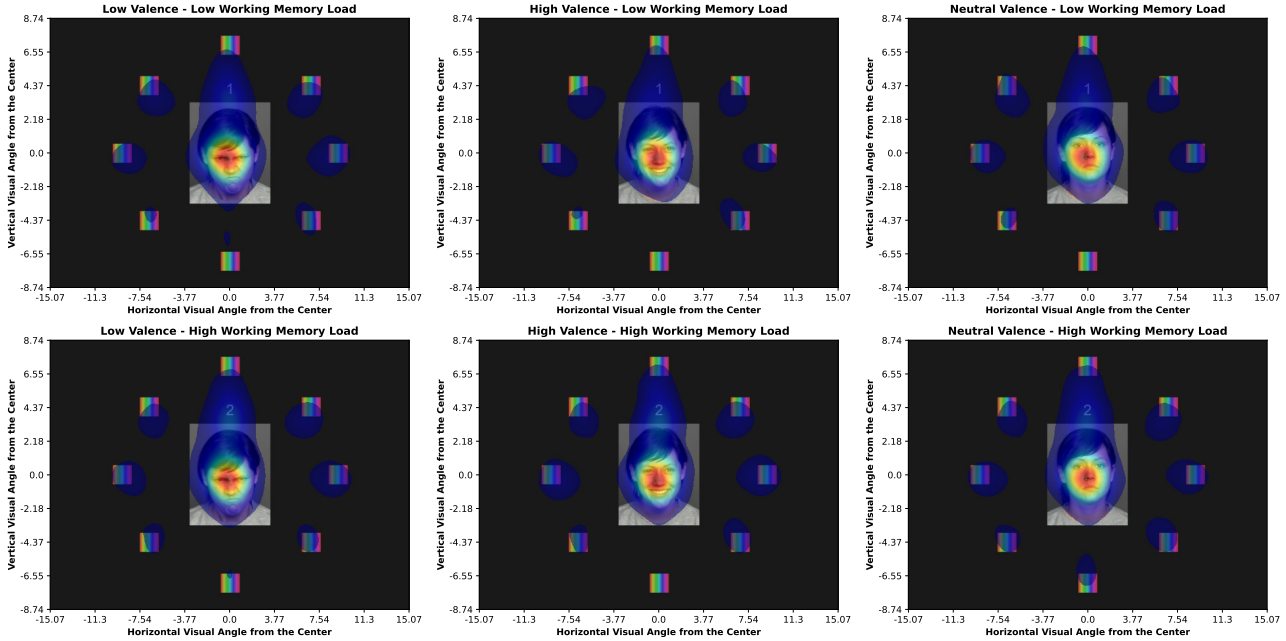
Descriptive Heatmap of the Fixations Averaged Across Subjects per Condition During the Facial Emotion Discrimination



Note. Colourmaps are normalised to a range between 0 and 1.

Supplementary Figure 6.9

Descriptive Heatmap of the Fixations Averaged Across Subjects per Condition During the N-Back Encoding/Retrieval Phase



Note. Colourmaps are normalised to a range between 0 and 1.

Supplementary References Study I

Bonnefond, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Current Biology*, *22*(20), 1969–1974. https://doi.org/10.1016/j.cub.2012.08.029

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, *23*(12), 1655–1665. https://doi.org/10.1038/s41593-020-00744-x

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. S. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267. https://doi.org/10.3389/fnins.2013.00267

Gyurkovics, M., Clements, G. M., Low, K. A., Fabiani, M., & Gratton, G. (2021b). The impact of 1/f activity and baseline correction on the results and interpretation of time-frequency analyses of EEG/MEG data: A cautionary tale. *NeuroImage*, *237*, 118192. https://doi.org/10.1016/j.neuroimage.2021.118192

Lin, F.-H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*, *31*(1), 160–171. https://doi.org/10.1016/j.neuroimage.2005.11.054

Proskovec, A. L., Wiesman, A. I., Heinrichs-Graham, E., & Wilson, T. W. (2019). Load effects on spatial working memory performance are linked to distributed alpha and beta oscillations. *Human Brain Mapping*, *40*(12), 3682–3689. https://doi.org/10.1002/hbm.24625

7 Supplementary Material for Study II

GAUDIE: Development, Validation, and Exploration of a Naturalistic German AUDItory Emotional Database

Katharina Lingelbach[1,2], Mathias Vukelić[2], and Jochem W. Rieger[1]

[1] Department of Psychology, University of Oldenburg, Oldenburg, Germany

[2] Fraunhofer Institute for Industrial Engineering IAO, Nobelstraße 12, 70569 Stuttgart, Germany

Supplementary Material to the article published as:

Lingelbach, K., Vukelić, M., & Rieger, J. W. (2024). GAUDIE: Development, validation, and exploration of a naturalistic German auditory emotional database. *Behavior Research Methods, 56*, 2049–2063. https://doi.org/10.3758/s13428-023-02135-z Supplementary material available at: https://osf.io/xyr6j/

7.1 Detailed Information on Audio Sequences

Supplementary Table 7.1

Detailed Information of the GAUDIE Database

ID	Condition	Title	Year	Producer	Distributor	Length in sec	Description of the Content
negative 1	negative	This is us - Season 1 Episode 18 *Moonshadow*; EIDR ID: 10.5240/DCCB-8610-6182-4E96-8E78-H	2017	Dan Fogelman, Jess Rosenthal, Donald Todd, Ken Olin, Charlie Gogolak, John Requa, Glenn Ficarra, Isaac Aptaker, Elizabeth Berger, Kay Oyegun	20th Television, NBC Universal Television Distribution	190	Television Series Scene: Argument between a couple (female and male) in which she accuses him of not supporting her in her career as a singer and he is disappointed that she is not grateful for the sacrifices he has made.
negative 2	negative	J'ai tué ma mère (engl.: I killed my mother); EIDR ID: 10.5240/8B5F-596B-30D3-B1D4-931B-N	2009	Xavier Dolan, Carole Mondello, Daniel Morin	Les cinémas nationaux de qualité K-Films Amérique	71	Movie Scene: Argument between teenager son and mother about going to a boarding school.
negative 3	negative	J'ai tué ma mère (engl.: I killed my mother); EIDR ID: 10.5240/8B5F-596B-30D3-B1D4-931B-N	2009	Xavier Dolan, Carole Mondello, Daniel Morin	Les cinémas nationaux de qualité K-Films Amérique	59	Movie Scene: Provocative conversation between a teenage son and his mother about inviting the son's friend to dinner.
negative 4	negative	J'ai tué ma mère (engl.: I killed my mother); EIDR ID: 10.5240/8B5F-596B-30D3-B1D4-931B-N	2009	Xavier Dolan, Carole Mondello, Daniel Morin	Les cinémas nationaux de qualité K-Films Amérique	230	Movie Scene: Argument in a car between mother and teenage son about her behaviour.
negative 5	negative	J'ai tué ma mère (engl.: I killed my mother); EIDR ID: 10.5240/8B5F-596B-30D3-B1D4-931B-N	2009	Xavier Dolan, Carole Mondello, Daniel Morin	Les cinémas nationaux de qualité K-Films Amérique	78	Movie Scene: Conversation between mother and teenage son about moving out and renting a flat for the son.
negative 6	negative	Marriage Story; EIDR ID: 10.5240/1B8A-F9DB-741E-513B-8FD8-Z	2019	David Heyman, Noah Baumbach	Netflix	393	Movie Scene: A couple (female and male) talk about their past and their feelings for each other. The talk ends up in a fight, and he wishes for her death.
negative 7	negative	Obsessed; EIDR ID: 10.5240/585F-803A-EC5A-4215-54F7-S	2008	Screen Gems, Rainforest Films, Parkwood Entertainment	Sony Pictures Releasing	146	Movie Scene: Argument between a couple (female and male) where she accuses him of having an affair and requires that he move out.
negative 8	negative	The Last Kiss; EIDR ID: 10.5240/14E9-0A1C-CBBD-E0AC-998F-P	2006	Gary Lucchesi, Tom Rosenberg, Marcus Viscidi, Andre Lamal	DreamWorks Pictures, Paramount Pictures	134	Movie Scene: Argument between a couple (female and male) where she accuses him of having an affair.

Continued on next page.

Supplementary Table 7.1 ***(continued)***

ID	Condition	Title	Year	Producer	Distributor	Length in sec	Description of the Content
negative 9	negative	Revolutionary Road; EIDR ID: 10.5240/0FD9-2D56-10BA-5DAD-3E41-F	2008	Bobby Cohen, Sam Mendes, Scott Rudin, John Hart	Paramount Vantage, Paramount Pictures	153	Movie Scene: Argument between a couple (female and male) about her intended abortion of their third child.
negative 10	negative	Revolutionary Road; EIDR ID: 10.5240/0FD9-2D56-10BA-5DAD-3E41-F	2008	Bobby Cohen, Sam Mendes, Scott Rudin, John Hart	Paramount Vantage, Paramount Pictures	134	Movie Scene: Argument in a car between a couple (female and male) where he first tries to conciliate with his wife but then gets angry.
negative 11	negative	Revolutionary Road; EIDR ID: 10.5240/0FD9-2D56-10BA-5DAD-3E41-F	2008	Bobby Cohen, Sam Mendes, Scott Rudin, John Hart	Paramount Vantage, Paramount Pictures	129	Movie Scene: Intense argument between a couple (female and male) about their relationship and feelings for each other.
neutral 1	neutral	Wetter heute - Weather Forecast 07.07.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by male speaker
neutral 2	neutral	Wetter heute - Weather Forecast 09.02.2020; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2020	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by female speaker
neutral 3	neutral	Wetter heute - Weather Forecast 10.05.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	138	Weather forecast by female speaker
neutral 4	neutral	Wetter heute - Weather Forecast 06.08.2019; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2019	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by female speaker
neutral 5	neutral	Wetter heute - Weather Forecast 21.07.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by male speaker
neutral 6	neutral	Wetter heute - Weather Forecast 08.05.2020; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2020	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by male speaker
neutral 7	neutral	Wetter heute - Weather Forecast 16.05.2020; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2020	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by male speaker
neutral 8	neutral	Wetter heute - Weather Forecast 02.07.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	117	Weather forecast by male speaker

Continued on next page.

Supplementary Table 7.1 ***(continued)***

ID	Condition	Title	Year	Producer	Distributor	Length in sec	Description of the Content
neutral 9	neutral	Wetter heute - Weather Forecast 01.07.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	118	Weather forecast by male speaker
neutral 10	neutral	Wetter heute - Weather Forecast 26.06.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by male speaker
neutral 11	neutral	Wetter heute - Weather Forecast 16.06.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by male speaker
neutral 12	neutral	Wetter heute - Weather Forecast 27.05.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by female speaker
neutral 13	neutral	Wetter heute - Weather Forecast 03.05.2021; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2021	wetter.com GmbH	wetter.com GmbH	120	Weather forecast by female speaker
neutral 14	neutral	Wetter heute - Weather Forecast 18.08.2020; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2020	wetter.com GmbH	wetter.com GmbH	150	Weather forecast by female speaker
neutral 15	neutral	Wetter heute - Weather Forecast 06.08.2020; EIDR ID: 10.5240/36A9-504A-2CF5-7258-45BA-8	2020	wetter.com GmbH	wetter.com GmbH	126	Weather forecast by female speaker
positive 1	positive	Comedy about romantic relationships - Part 1	2017	Hazel Brugger	Agentur ZweiR	136	Comedian show: Female comedian talks about relationship arguments from the female perspective.
positive 2	positive	Comedy about romantic relationships - Part 2	2017	Hazel Brugger	Agentur ZweiR	114	Comedian show: Female comedian talks about relationship arguments from the female perspective.
positive 3	positive	Känguru Chroniken - Das Känguru-Manifest - Chapter 42 *Bier*	2011	Marc-Uwe Kling	Verlag Voland & Quist GmbH	223	Audiobook: A woman and a little boy are suddenly standing in the flat to talk to the male narrator and a kangaroo about religion and God.
positive 4	positive	Känguru Chroniken – Das Känguru-Manifest – Chapter 13 *Street Art*	2011	Marc-Uwe Kling	Verlag Voland & Quist GmbH	138	Audiobook: The male narrator and a Kangaroo invent socially critical misattributed quotes.

Continued on next page.

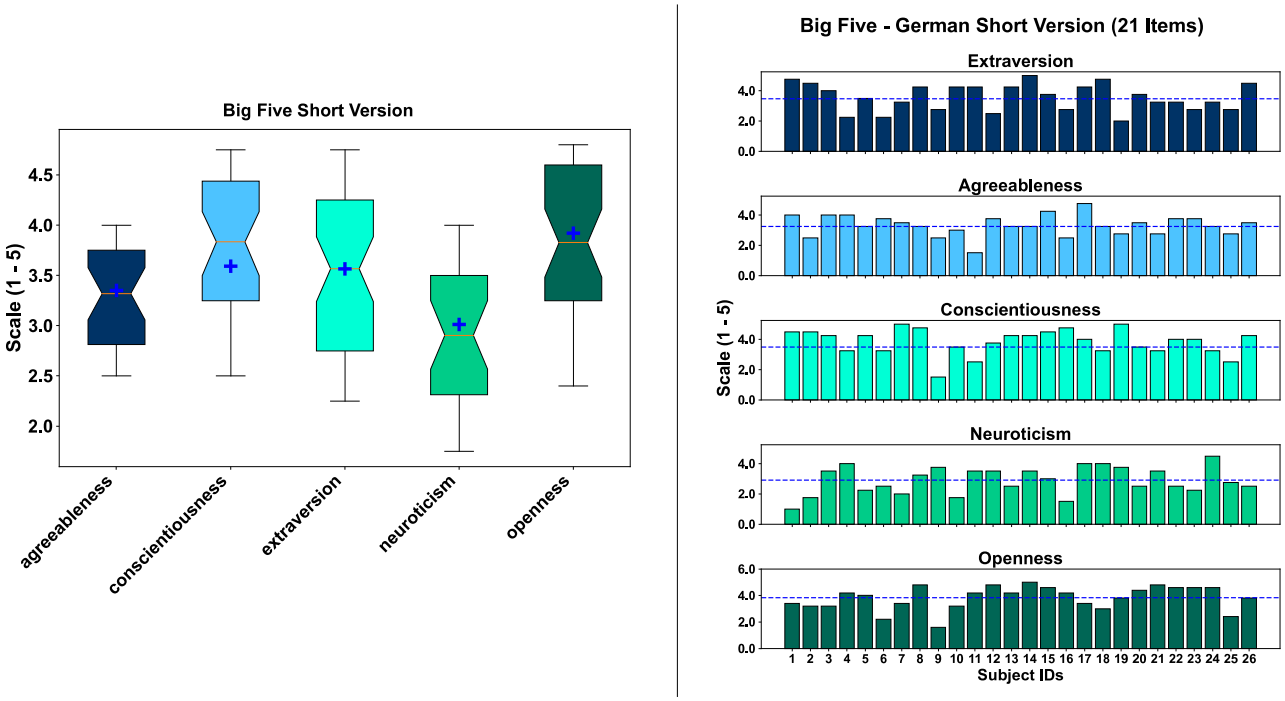
Supplementary Table 7.1 ***(continued)***

ID	Condition	Title	Year	Producer	Distributor	Length in sec	Description of the Content
positive 5	positive	Känguru Chroniken – Das Känguru-Manifest – Chapter 14 *Jörn Dwigs und der Deutsche Buchpreis der Ullstein Buchverlage*	2011	Marc-Uwe Kling	Verlag Voland & Quist GmbH	127	Audiobook: The male narrator and a kangaroo attend an award ceremony for a literary prize.
positive 6	positive	Känguru Chroniken – Das Känguru-Manifest – Chapter 11 *Vier Fäuste außer Rand und Band*	2011	Marc-Uwe Kling	Verlag Voland & Quist GmbH	109	Audiobook: A male narrator and the kangaroo are in different scenes that allude to the famous film ”Vier Fäuste außer Rand und Band”.
positive 7	positive	Comedy about comedy	2021	Moritz Neumeier	Agentur Oschatz Management,	148	Comedian show: The male comedian talks about the comedy in Germany.
positive 8	positive	Comedy about politically correct language	2021	Moritz Neumeier	Agentur Oschatz Management,	206	Comedian show: The male comedian talks about the use of politically correct language.
positive 9	positive	Comedy about weddings	2017	Nico Semsrott	Office of MEP Nico Semsrott	217	Comedian show: The male comedian talks about weddings.
positive 10	positive	Comedy about bodily reaction during anger	2011	Otto Waalkes	Otto Waalkes	185	Comedian show: The male comedian demonstrates a scene of anger from the body parts' point of view.
positive 11	positive	Comedy about flirt tips	2017	Till Reiners	Agentur Oschatz Management	336	Comedian show: The male comedian advises on how to flirt.

7.2 Sample Characteristics: Personality and Well-Being

Supplementary Figure 7.1

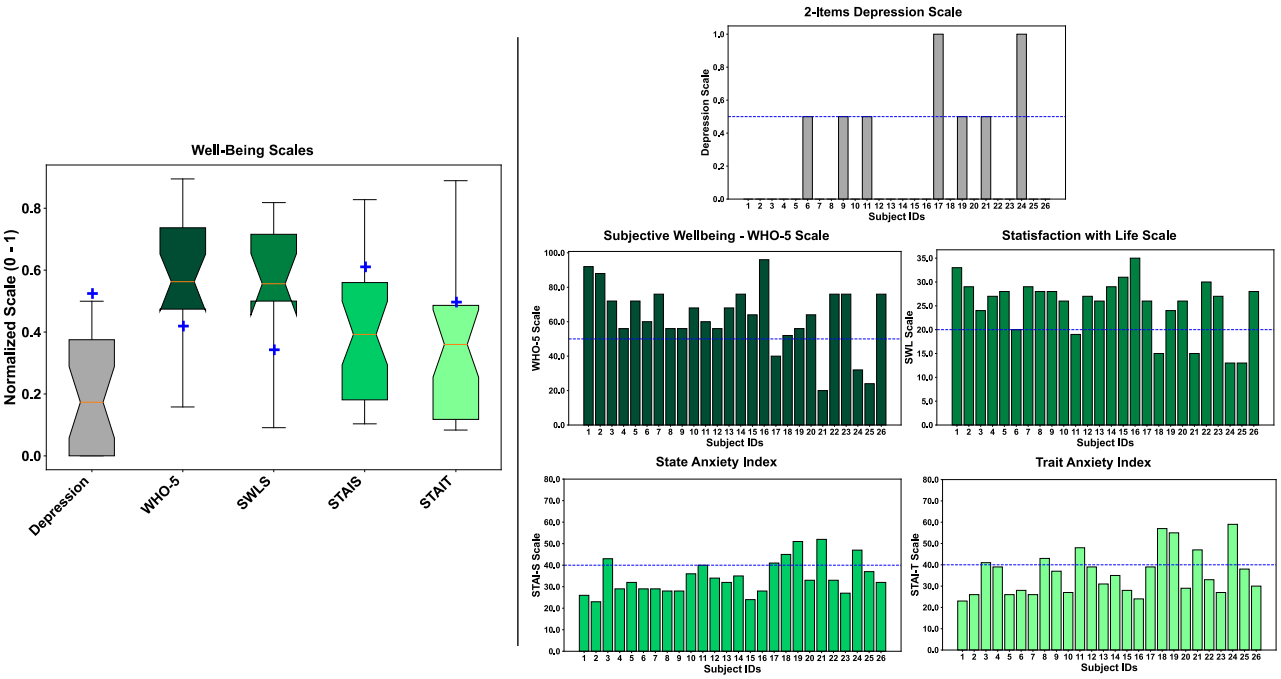
Results of the Personality Questionnaire Big Five - German Short Version With 21 Items (Rammstedt & John, 2005)



Note. Left: Bootstrapped mean and 2.5th and 97.5th confidence intervals (CIs) over participants for each scale of the Big Five indicated by notches of the boxes. The orange line represents the bootstrapped mean. Boxes comprise $50\ \%$ of the distribution from the 25th to the 75th quartile, while ends of the whiskers represent the 5th and 95th quartiles of the distribution. Right: Individual scores of the Big Five scales. The cross and dashed lines represent the previously published reference values from 850 young participants, mainly from German universities. Extraversion: 3.47, Agreeableness: 3.26, Conscientiousness: 3.50, Neuroticism: 2.92; Openness to New Experiences: 3.83 (Rammstedt & John, 2005).

Supplementary Figure 7.2

Results in Questionnaires Associated With Subjective Well-Being (Janke & Glöckner-Rist, 2012; Whooley et al., 1997; World Health Organization, 1998), and Anxiety (Laux et al., 1981)



Note. Left: Bootstrapped mean and 2.5th and 97.5th confidence intervals (CIs) over participants for each scale indicated by the notches of the boxes. The orange line represents the bootstrapped mean. Boxes comprise $50\ \%$ of the distribution from the 25th to the 75th quartile, while ends of the whiskers represent the 5th and 95th quartiles of the distribution. Right: Individual scores of the scales. The cross and dashed lines represent a threshold for a clinically critical score. Depression: 0.5 (Whooley et al., 1997), WHO-5: 50 (World Health Organization, 1998), SWLS: 20 (Janke & Glöckner-Rist, 2012), STAI-S: 40, and STAI-T: 40 (Laux et al., 1981).

Supplementary Table 7.2

Bootstrapped Mean and 2.5th and 97.5th Confidence Interval (CI) Over Participants for Each Scale of the Big Five and the Questionnaires Related to Subjective Well-Being and Anxiety

Big Five				**Subjective Well-Being**			
	2.5th CI	***M***	***97.5th CI***		***2.5th CI***	***M***	***97.5th CI***
Agreeableness	3.05	3.32	3.57	Depression	0.06	0.17	0.31
Conscientiousness	3.51	3.84	4.14	WHO-5	55.54	62.80	69.85
Extraversion	3.23	3.57	3.89	SWLS	22.85	25.21	27.31
Neuroticism	2.56	2.91	3.25	STAIS	31.50	34.44	37.54
Openness to New Experiences	3.48	3.83	4.15	STAIT	32.15	35.97	40.00

Note. M: Mean; CI: Confidence Interval; WHO-5: World Health Organization Well-Being Index; SWLS: Satisfaction With Life Scale; STAIS: State Anxiety; STAIT: Trait Anxiety.

7.2.1 Internal Consistency of the Scales

Supplementary Table 7.3

Internal Consistency (Cronbach's α) and Bootstrapped 5th and 95th Confidence Intervals (CI) for the Used Personality, Well-Being, and Anxiety Scales

Scale	***Cronbach α***	***5th CI***	***95th CI***
BIG 5 Extraversion	0.842	0.712	0.922
BIG 5 Agreeableness	0.557	0.194	0.781
BIG 5 Consciousness	0.835	0.701	0.919
BIG 5 Neuroticism	0.738	0.523	0.870
BIG 5 Openness to New Experience	0.809	0.662	0.904
Depression	0.529	-0.050	0.789
WHO-5	0.836	0.710	0.918
SWLS	0.852	0.738	0.926
STAI S	0.849	0.750	0.922
STAI T	0.610	0.354	0.797

Note. Cronbach's α was calculated using the Python toolbox `pingouin` (version 0.4.0). CI: Confidence Interval; WHO-5: World Health Organization Well-Being Index; SWLS: Satisfaction With Life Scale; STAIS: State Anxiety; STAIT: Trait Anxiety.

7.3 Study Instruction

7.3.1 German Version

Lieber Studienteilnehmer*in,

in dieser Studie werden Sie Audiosequenzen anhören und gebeten, diese auf unterschiedlichen Skalen zu bewerten. Dabei interessiert uns, die in Ihnen ausgelösten Empfindungen - *nicht* die Inhalte der Szene oder was die Protagonist*innen empfinden. Sie werden die Sequenzen anhand folgender Dimensionen bewerten:

- **Valenz**, d. h. emotionale Wertigkeit - ob die Audiosequenz bei Ihnen ein positives oder negatives Gefühl hervorgerufen hat,
- **Grad der Aktivität**, d. h. ob Sie die Audiosequenz als beruhigend oder aktivierend/erregend wahrgenommen haben, und
- **Grad der Dominanz und Kontrolle**, d. h. ob die Audiosequenz bei Ihnen das Gefühl von Aufdringlichkeit und Beherrschung oder Selbstkontrolle und wahrgenommene Zurückhaltung ausgelöst hat.

Zudem fragen wir Sie nach den von Ihnen erlebten Emotionen und ob Ihnen die Audiosequenz vertraut ist. Während der Präsentation der Audiosequenz geben Sie bitte kontinuierlich über das Verstellen des Schiebereglers eine Bewertung ab.

7.3.2 English Translation

Dear Participant,

In this study, you will listen to audio sequences and be asked to rate them on different scales. We are interested in the sensations and feelings triggered in you – *not* the content of the scene or what the protagonists might experience. You will rate the sequences on the following dimensions:

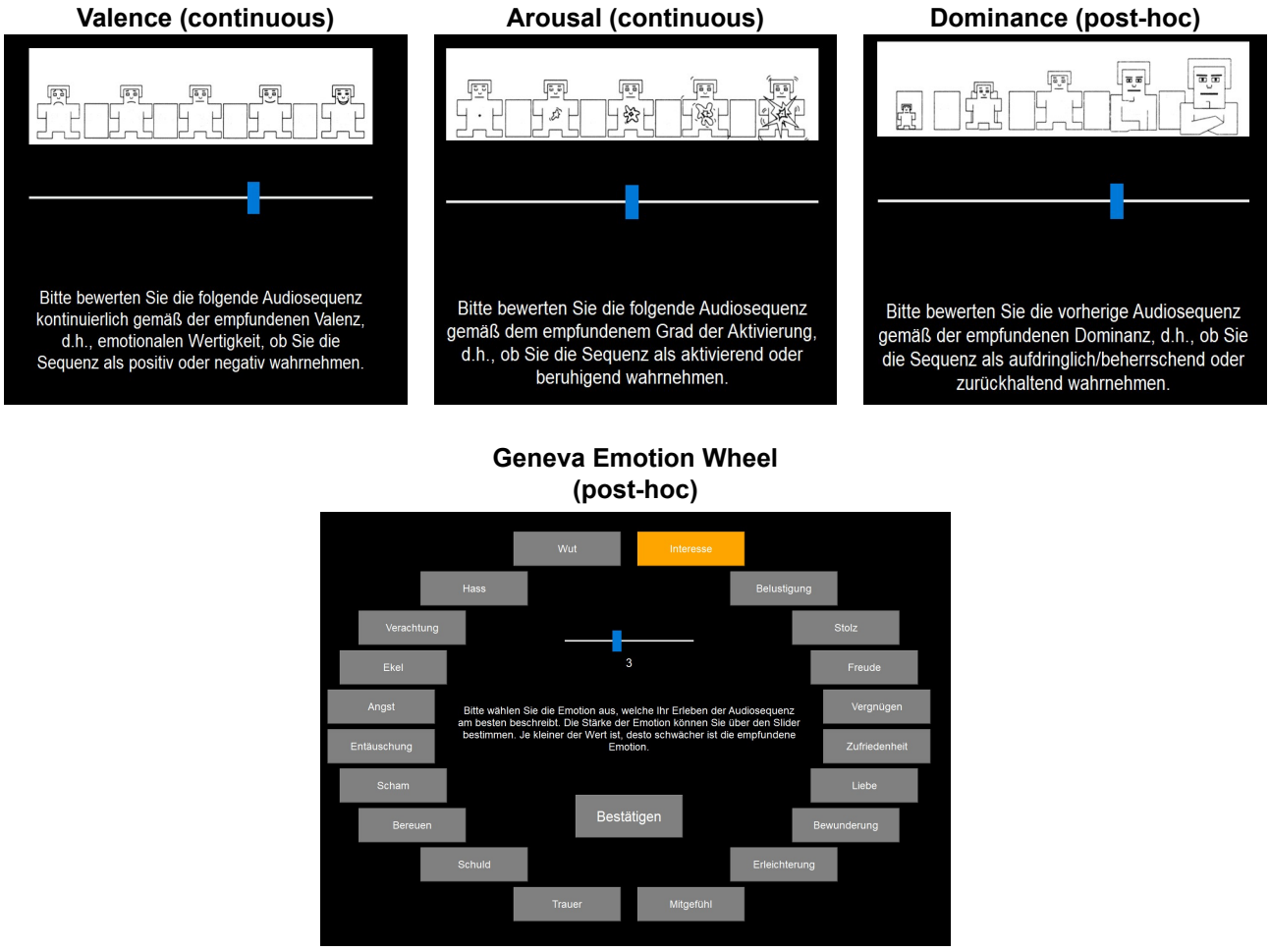
- **Valence**, i.e., emotional valence – whether the audio sequence evoked a positive or negative feeling in you,
- **Arousal**, i.e., whether you perceived the audio sequence as calming or activating/exciting, and
- **Dominance and control**, i.e., whether the audio sequence elicited in you a feeling of being controlled and dominated or self-control and perceived restraint.

In addition, we ask you about the emotions you experienced and whether the audio sequence is familiar to you. During the presentation of the audio sequence, please provide a rating continuously by adjusting the slider.

7.4 Rating Scales Used in the Validation Study

Supplementary Figure 7.3

Illustration of the Presented SAM and GEW Scales



Note. The SAM scales range from 0 to 100 (left valence, middle arousal, and right dominance). The GEW strength scale ranges from 1 to 6 on the slider, and the highlighted field indicates the chosen prevailing emotion.

7.5 Supplementary Tables of the SAM Ratings

Supplementary Table 7.4

Bootstrapped Average Valence, Arousal, and Dominance Ratings per Audio Sequence

Negative Audio Sequences				**Neutral Audio Sequences**				**Positive Audio Sequences**			
					Valence						
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	20.37	24.43	28.59	1	50.59	53.84	57.03	1	68.95	73.34	77.65
2	31.14	34.57	38.34	2	44.20	50.23	56.73	2	59.32	66.72	73.39
3	24.20	30.37	36.45	3	51.56	57.30	63.45	3	67.24	74.89	81.93
4	28.34	31.29	34.30	4	44.36	47.91	51.52	4	66.97	76.07	83.59
5	31.76	36.61	40.73	5	52.50	57.32	61.86	5	51.92	62.20	70.66
6	17.27	20.74	24.01	6	51.12	55.04	59.10	6	66.73	73.72	80.33
7	10.72	15.05	19.77	7	48.90	53.16	57.76	7	51.29	59.93	67.36
8	13.36	16.61	20.44	8	44.35	53.20	60.32	8	52.78	61.30	69.01
9	17.13	21.85	26.67	9	43.01	48.26	53.48	9	56.49	66.46	76.31
10	19.45	23.62	27.67	10	52.07	56.57	60.30	10	59.71	69.04	77.72
11	19.82	26.70	34.04	11	47.27	52.89	58.42	11	67.59	72.89	78.21
				12	37.92	44.18	49.92				
				13	37.26	46.28	54.38				
				14	42.86	47.56	51.83				
				15	48.54	51.58	54.88				
Average	21.23	25.62	30.09		46.43	51.69	56.73		60.82	68.78	76.01

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. *Continued on next page.*

Supplementary Table 7.4 ***(continued)***

Bootstrapped Average Valence, Arousal, and Dominance Ratings per Audio Sequence

Negative Audio Sequences				**Neutral Audio Sequences**				**Positive Audio Sequences**			
					Arousal						
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	65.53	74.50	83.26	1	19.93	29.44	39.88	1	42.39	53.30	64.33
2	59.22	65.91	72.96	2	21.94	32.75	44.22	2	49.32	56.43	63.65
3	58.11	66.63	75.02	3	22.66	32.26	42.46	3	57.67	65.59	74.00
4	59.22	66.48	74.25	4	21.79	32.32	43.27	4	45.59	59.10	71.52
5	52.58	58.86	65.04	5	29.67	38.07	46.89	5	50.32	58.48	66.24
6	65.72	74.14	82.87	6	29.33	37.93	46.97	6	40.74	52.91	63.98
7	69.38	75.83	82.69	7	15.05	25.50	36.46	7	48.41	56.54	63.75
8	73.58	78.57	82.39	8	23.76	30.21	36.49	8	38.76	53.63	67.45
9	67.50	73.34	79.46	9	18.24	26.19	34.71	9	52.29	60.06	67.74
10	62.31	67.97	74.53	10	21.10	31.32	43.05	10	45.41	56.54	66.52
11	67.07	74.15	80.83	11	26.23	35.18	44.19	11	45.97	54.88	62.63
				12	18.63	26.44	34.85				
				13	20.00	28.48	37.22				
				14	18.00	26.89	36.00				
				15	23.62	32.60	41.01				
Average	63.66	70.58	77.57		22.00	31.04	40.51		46.99	57.04	66.53

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. *Continued on next page.*

Supplementary Table 7.4 ***(continued)***

Bootstrapped Average Valence, Arousal, and Dominance Ratings per Audio Sequence

Negative Audio Sequences				**Neutral Audio Sequences**				**Positive Audio Sequences**			
				Dominance							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	77.46	86.60	94.69	1	21.69	30.65	39.46	1	49.38	57.12	64.85
2	73.38	82.35	90.92	2	23.23	34.75	47.38	2	43.85	55.46	66.00
3	77.85	85.23	92.31	3	21.77	29.24	37.15	3	53.62	59.86	66.15
4	81.69	88.88	95.08	4	13.77	22.48	31.69	4	42.69	54.76	65.38
5	76.38	84.46	91.15	5	19.38	28.09	37.39	5	45.92	54.59	61.92
6	84.00	90.45	96.08	6	21.54	35.48	50.31	6	41.31	51.04	59.46
7	80.54	86.67	91.92	7	19.23	28.08	36.69	7	53.62	63.52	72.92
8	77.84	86.02	92.92	8	22.23	33.80	47.69	8	42.23	53.95	65.62
9	76.46	84.35	91.08	9	13.62	26.79	42.15	9	46.31	57.74	68.62
10	78.61	85.14	91.46	10	26.46	39.57	53.77	10	46.77	58.32	68.23
11	79.54	85.48	90.92	11	21.61	29.97	38.54	11	46.46	56.15	63.77
				12	21.77	27.52	33.15				
				13	23.31	33.93	46.00				
				14	14.62	24.16	34.23				
				15	19.92	26.68	33.77				
Average	78.52	85.96	92.59		20.28	30.08	40.63		46.56	56.59	65.72

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile.

Supplementary Table 7.5

Bootstrapped Average Standard Deviation of the Valence and Arousal Ratings per Audio Sequence

Negative Audio Sequences				Neutral Audio Sequences				Positive Audio Sequences			
				Valence Standard Deviation							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	6.63	9.34	12.22	1	5.59	9.34	13.26	1	7.28	9.94	12.89
2	6.19	8.69	11.86	2	3.00	5.24	7.53	2	5.46	8.36	11.39
3	2.11	5.09	8.64	3	3.45	5.66	7.92	3	3.82	6.17	8.69
4	11.83	14.03	16.20	4	3.45	7.48	11.82	4	4.97	7.89	10.73
5	10.23	13.58	16.66	5	2.86	5.48	8.93	5	5.59	7.46	9.32
6	10.12	11.70	13.25	6	7.22	11.46	16.22	6	4.97	7.67	10.63
7	3.14	4.53	6.02	7	3.31	6.40	10.03	7	4.27	8.56	13.26
8	4.77	6.81	8.96	8	4.58	7.40	10.07	8	6.50	8.91	11.57
9	4.90	6.96	9.15	9	4.34	8.01	11.92	9	5.27	8.31	11.49
10	4.65	6.31	7.87	10	3.90	7.73	12.12	10	6.32	8.07	9.63
11	7.52	10.80	14.13	11	6.02	12.23	18.72	11	5.17	6.71	8.09
				12	3.20	6.23	9.62				
				13	4.48	7.30	10.47				
				14	2.44	4.76	7.43				
				15	4.43	6.62	8.88				
Average	6.55	8.89	11.36		4.15	7.42	11.00		5.42	8.00	10.70

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. *Continued on next page.*

Supplementary Table 7.5 ***(continued)***

Bootstrapped Average Standard Deviation of the Valence and Arousal Ratings per Audio Sequence

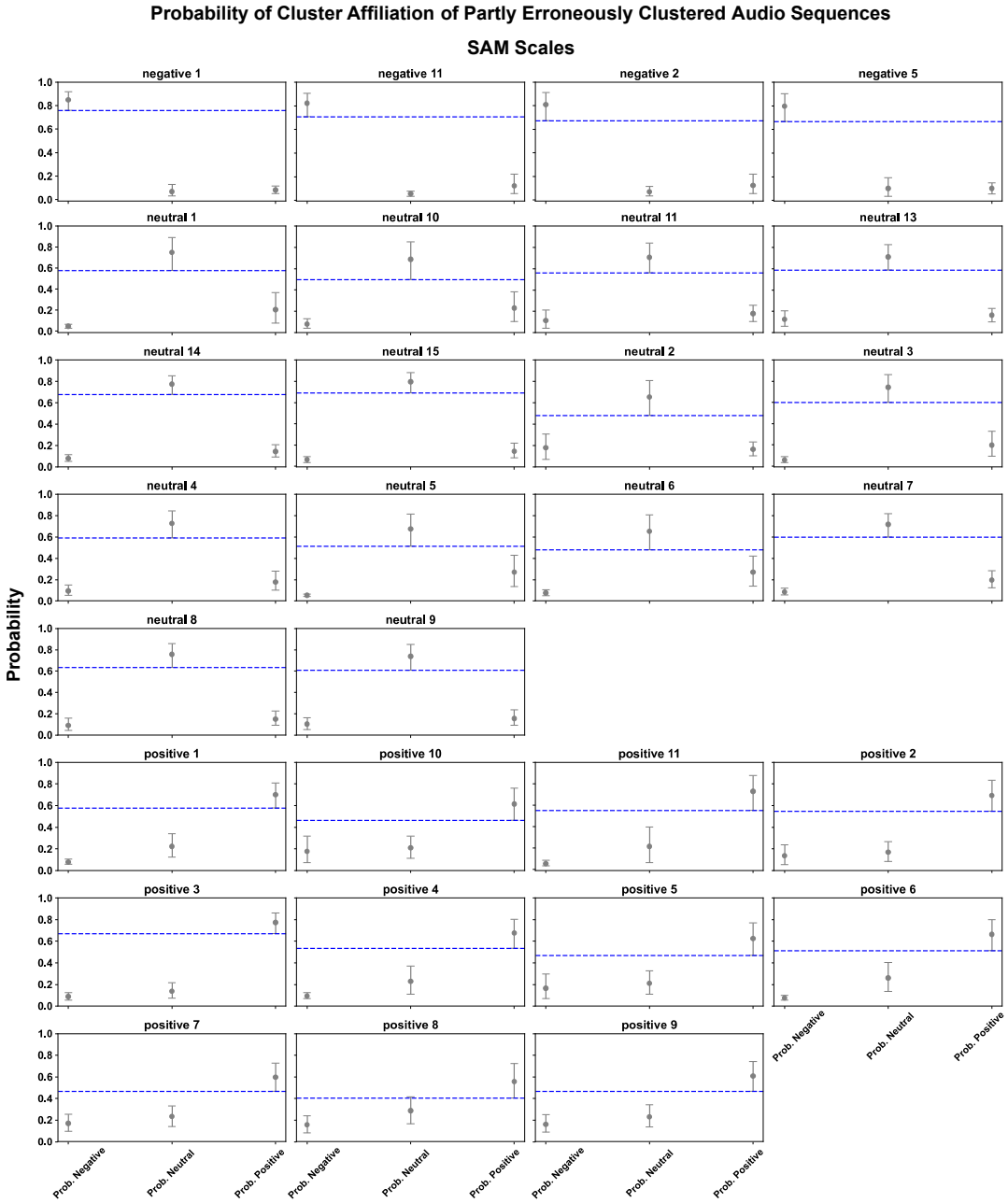
Negative Audio Sequences				**Neutral Audio Sequences**				**Positive Audio Sequences**			
				Arousal Standard Deviation							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	7.43	12.14	17.02	1	5.98	10.63	15.47	1	5.30	7.68	10.16
2	5.07	9.54	15.32	2	1.59	5.12	9.16	2	2.91	5.77	8.75
3	2.75	6.25	10.97	3	2.00	4.21	6.72	3	4.73	8.95	14.25
4	13.72	18.92	23.64	4	2.41	5.70	9.49	4	5.56	8.24	10.65
5	9.51	14.91	20.20	5	3.57	6.90	10.91	5	5.83	8.49	11.31
6	7.83	13.79	19.90	6	3.22	7.28	11.84	6	3.33	5.60	8.12
7	3.20	5.97	8.68	7	3.28	6.55	10.35	7	2.52	6.16	10.21
8	6.17	9.39	12.71	8	3.67	7.04	10.85	8	4.68	7.87	11.53
9	5.27	8.38	11.46	9	2.94	5.91	9.49	9	4.38	7.62	10.97
10	9.04	12.65	15.90	10	2.43	5.98	10.24	10	5.36	7.86	10.17
11	9.09	13.57	18.16	11	6.66	13.52	20.45	11	6.09	8.82	11.65
				12	1.29	3.11	5.33				
				13	4.03	7.67	11.39				
				14	2.73	4.79	7.12				
				15	3.27	6.05	9.01				
Average	7.19	11.41	15.81		3.27	6.70	10.52		4.61	7.55	10.71

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile.

7.6 Probabilities of Cluster Affiliation

Supplementary Figure 7.4

Bootstrapped Mean and 2.5th and 97.5th Confidence Interval (CI) of the Probability of Cluster Affiliation per Cluster for Critical Audio Sequences



Note. Critical audio sequences are audio sequences which were at least for one participant clustered into the wrong (i.e., not the hypothetical) emotion condition. Dashed line represents the lower confidence interval of the hypothetically correct cluster. The clustering is based on the three SAM rating scales as feature set. Prob: Probability to belong to the cluster.

Supplementary Table 7.6

Probabilities and Bootstrapped Confidence Intervals for Assignment to Negative, Neutral, and Positive Clusters per Sequence

	Probability Negative Cluster			**Probability Neutral Cluster**			**Probability Positive Cluster**		
ID	*2.5th CI*	*M*	*97.5th CI*	*2.5th CI*	*M*	*97.5th CI*	*2.5th CI*	*M*	*97.5th CI*
negative 1	0.76	0.85	0.92	0.03	0.07	0.13	0.05	0.08	0.12
negative 2	0.67	0.81	0.91	0.03	0.07	0.11	0.06	0.12	0.22
negative 5	0.67	0.80	0.90	0.04	0.10	0.19	0.06	0.10	0.15
negative 11	0.71	0.82	0.90	0.04	0.06	0.08	0.06	0.12	0.22
neutral 1	0.03	0.05	0.07	0.58	0.75	0.89	0.08	0.21	0.37
neutral 2	0.07	0.18	0.31	0.48	0.66	0.81	0.11	0.17	0.23
neutral 3	0.03	0.06	0.09	0.60	0.74	0.86	0.09	0.20	0.33
neutral 4	0.05	0.09	0.15	0.59	0.73	0.84	0.10	0.18	0.28
neutral 5	0.04	0.05	0.07	0.51	0.68	0.82	0.14	0.27	0.43
neutral 6	0.05	0.08	0.11	0.48	0.65	0.81	0.14	0.27	0.42
neutral 7	0.06	0.09	0.12	0.60	0.72	0.82	0.12	0.20	0.28
neutral 8	0.04	0.09	0.16	0.64	0.76	0.86	0.09	0.15	0.23
neutral 9	0.05	0.10	0.16	0.61	0.74	0.85	0.09	0.16	0.24
neutral 10	0.04	0.08	0.13	0.50	0.69	0.85	0.10	0.23	0.38
neutral 11	0.04	0.11	0.21	0.56	0.71	0.84	0.11	0.18	0.26
neutral 13	0.06	0.12	0.21	0.59	0.71	0.82	0.10	0.17	0.23
neutral 14	0.05	0.08	0.11	0.68	0.78	0.85	0.09	0.14	0.21
neutral 15	0.04	0.06	0.09	0.69	0.80	0.88	0.08	0.14	0.22

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. *Continued on next page.*

Supplementary Table 7.6 ***(continued)***

Probabilities and Bootstrapped Confidence Intervals for Assignment to Negative, Neutral, and Positive Clusters per Sequence

	Probability Negative Cluster			**Probability Neutral Cluster**			**Probability Positive Cluster**		
ID	*2.5th CI*	*M*	*97.5th CI*	*2.5th CI*	*M*	*97.5th CI*	*2.5th CI*	*M*	*97.5th CI*
positive 1	0.06	0.08	0.11	0.12	0.22	0.34	0.58	0.70	0.81
positive 2	0.05	0.14	0.24	0.09	0.17	0.27	0.55	0.69	0.84
positive 3	0.06	0.09	0.12	0.07	0.14	0.22	0.67	0.77	0.86
positive 4	0.07	0.09	0.12	0.11	0.23	0.37	0.53	0.68	0.80
positive 5	0.07	0.17	0.30	0.11	0.21	0.33	0.47	0.62	0.77
positive 6	0.05	0.08	0.10	0.14	0.26	0.40	0.51	0.66	0.80
positive 7	0.10	0.17	0.26	0.14	0.23	0.33	0.47	0.60	0.73
positive 8	0.08	0.16	0.24	0.17	0.29	0.42	0.40	0.56	0.72
positive 9	0.09	0.16	0.25	0.14	0.23	0.34	0.47	0.61	0.74
positive 10	0.07	0.18	0.32	0.11	0.21	0.32	0.47	0.61	0.76
positive 11	0.03	0.06	0.09	0.07	0.22	0.40	0.55	0.73	0.88

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile.

7.7 Familiarity Ratings

Supplementary Table 7.7

Bootstrapped Average Familiarity Ratings per Audio Sequence

Negative Audio Sequences				**Neutral Audio Sequences**				**Positive Audio Sequences**			
				Familiarity							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	4.54	20.43	39.08	1	25.15	45.89	67.69	1	24.69	42.54	61.92
2	10.85	22.75	36.38	2	40.00	55.16	69.77	2	18.61	37.12	56.69
3	4.46	14.40	26.92	3	29.31	48.19	66.77	3	31.69	51.77	71.54
4	10.69	19.30	29.00	4	31.69	51.72	71.31	4	42.61	62.21	81.00
5	8.92	20.01	33.00	5	35.92	57.73	78.54	5	27.85	49.62	70.93
6	8.77	27.63	49.31	6	45.85	58.36	69.92	6	17.77	33.14	50.46
7	5.38	21.47	40.08	7	36.46	53.93	71.23	7	12.38	26.72	43.62
8	8.54	16.27	24.92	8	37.15	54.54	71.08	8	12.23	32.18	54.77
9	9.92	24.40	42.08	9	30.77	51.35	70.85	9	26.77	46.27	66.23
10	9.00	19.47	31.08	10	24.00	45.28	68.16	10	22.38	42.08	63.31
11	15.69	29.74	45.00	11	24.08	41.43	60.16	11	20.23	37.62	56.15
				12	32.31	51.18	70.77				
				13	28.69	46.31	64.08				
				14	27.62	47.25	67.23				
				15	39.38	56.18	73.31				
Average	8.80	21.44	36.08		32.56	50.97	69.39		23.38	41.93	61.51

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile.

Supplementary Table 7.8

Pearson Correlations of the SAM Ratings Valence, Arousal, and Dominance With Familiarity

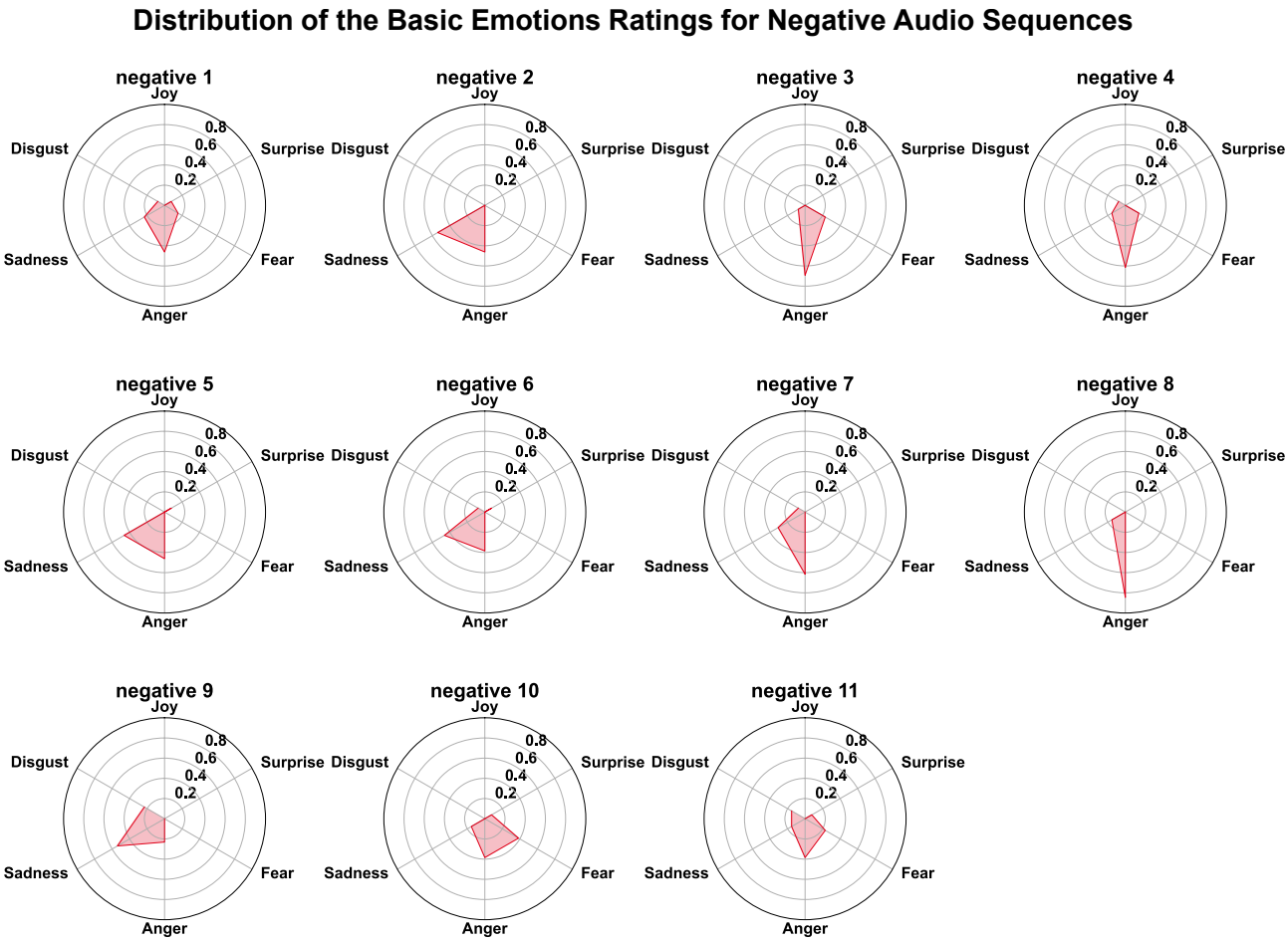
Condition	Correlation Pairs	*Pearson's r*	*p*	*corrected p*
negative	Arousal – Familiarity	-.042	.617	.730
	Valence – Familiarity	.070	.406	.730
	Dominance – Familiarity	.046	.586	.730
neutral	Arousal – Familiarity	.153	.032	.146
	Valence – Familiarity	-.008	.907	.907
	Dominance – Familiarity	.100	.163	.488
positive	Arousal – Familiarity	-.048	.566	.730
	Valence – Familiarity	.219	.008	.076
	Dominance – Familiarity	.038	.649	.730

Note. Obtained *p*-values are corrected for multiple testing using the False Discovery Rate method.

7.8 Discrete Emotion Ratings

Supplementary Figure 7.5

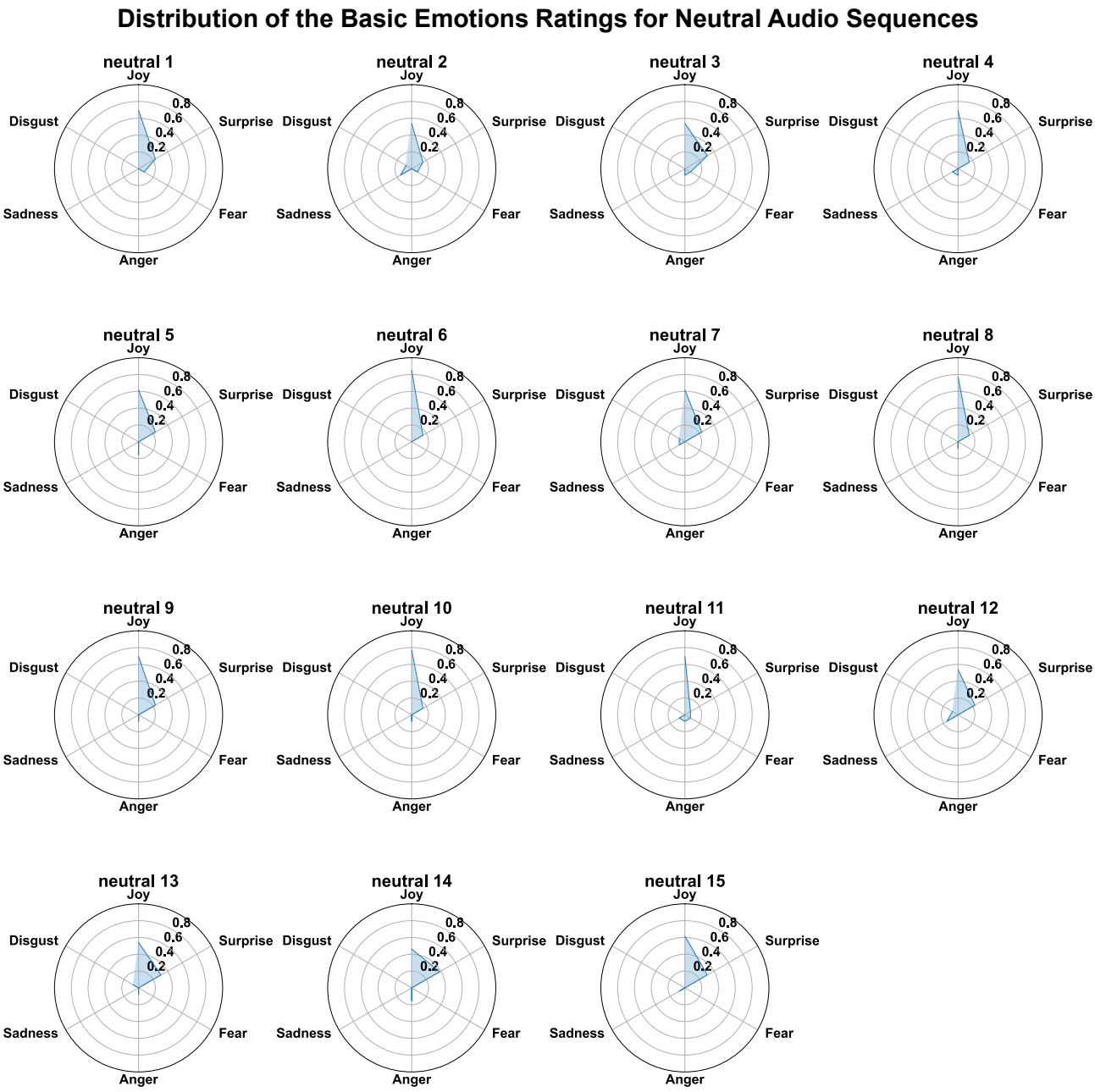
Frequencies of the Basic Emotions Ratings for the Negative Audio Sequences



Note. Frequencies are normalised between 0 and 1 to represent relative frequencies.

Supplementary Figure 7.6

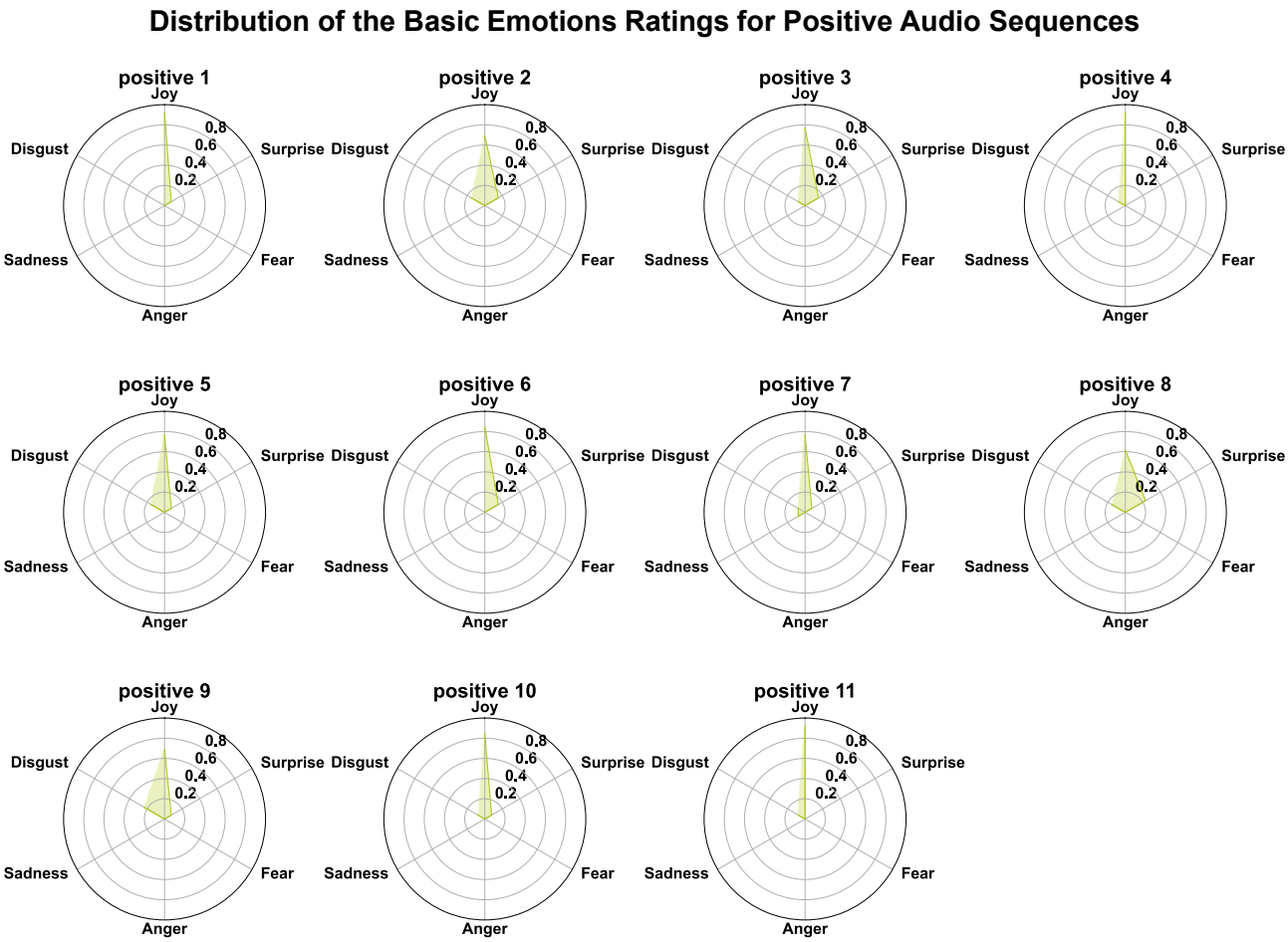
Frequencies of the Basic Emotions Ratings for the Neutral Audio Sequences



Note. Frequencies are normalised between 0 and 1 to represent relative frequencies.

Supplementary Figure 7.7

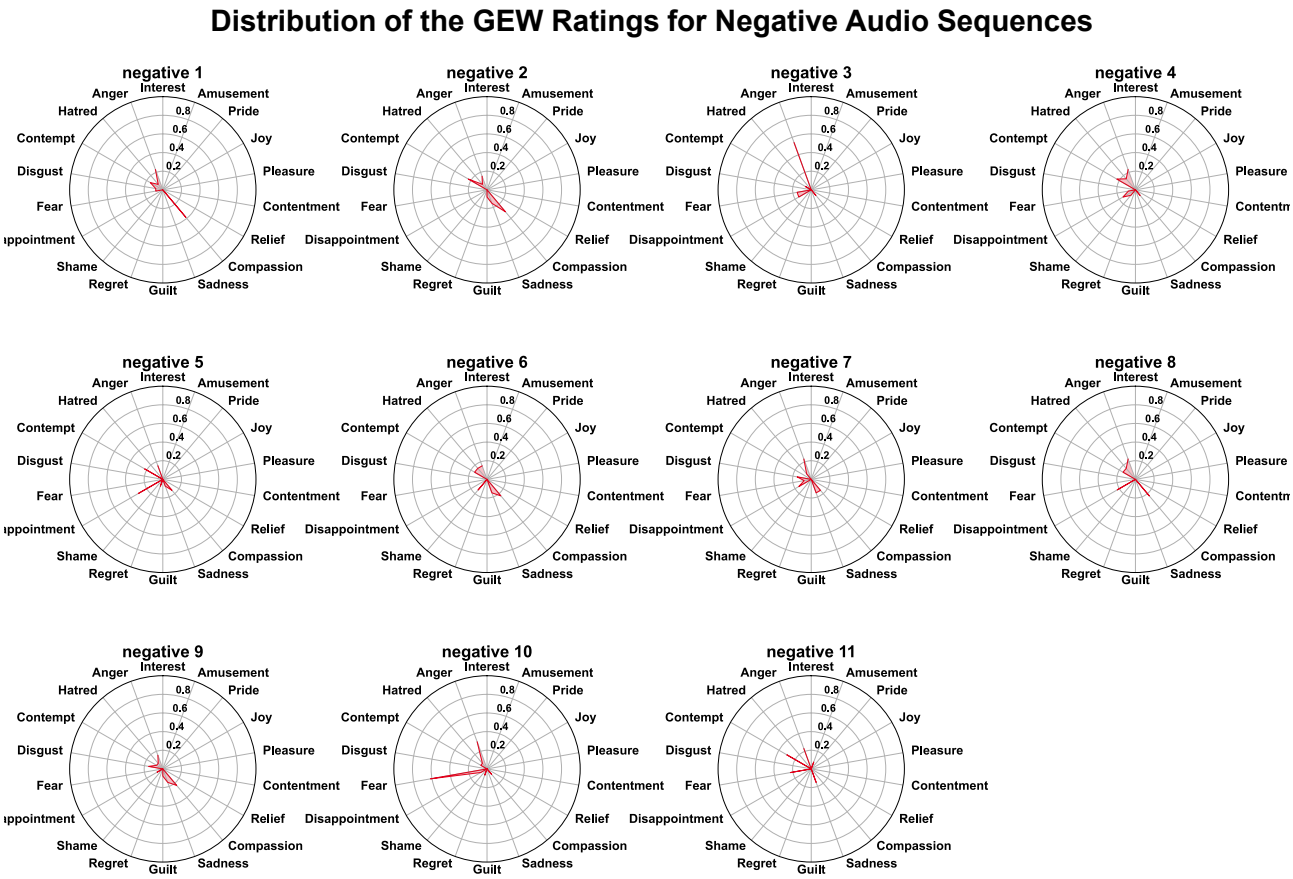
Frequencies of the Basic Emotions Ratings for the Positive Audio Sequences



Note. Frequencies are normalised between 0 and 1 to represent relative frequencies.

Supplementary Figure 7.8

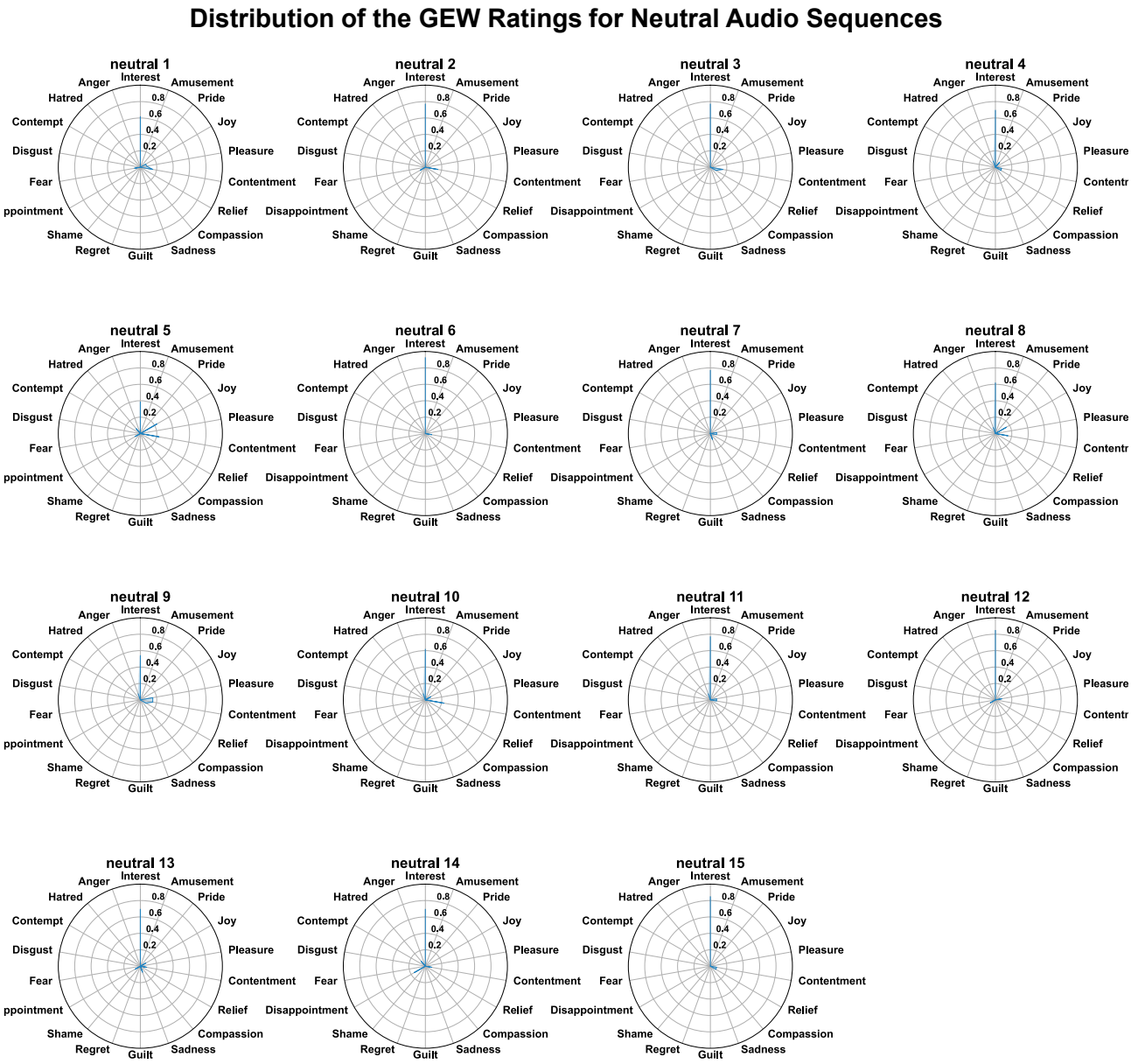
Frequencies of the GEW Emotions Ratings for the Negative Audio Sequences



Note. Frequencies are normalised between 0 and 1 to represent relative frequencies.

Supplementary Figure 7.9

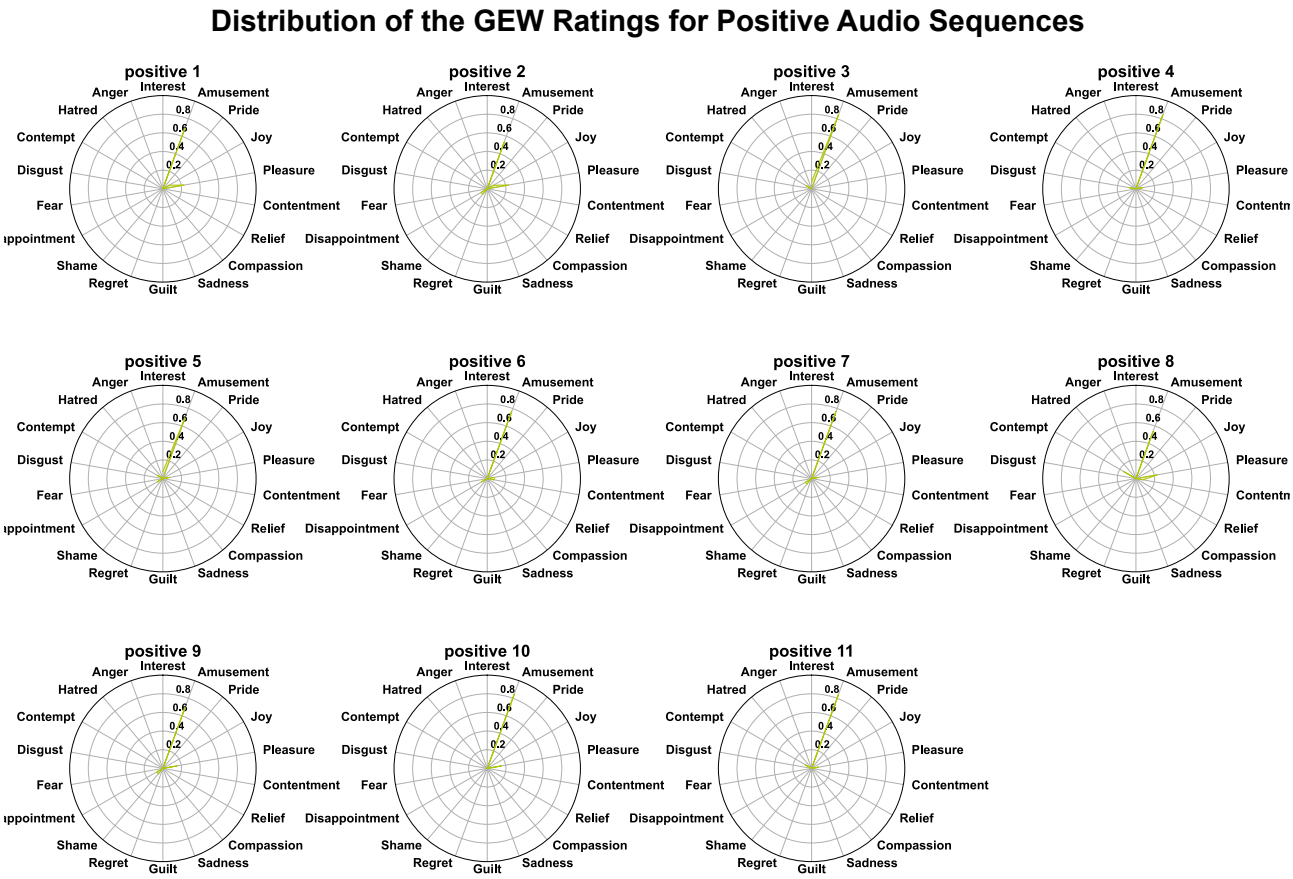
Frequencies of the GEW Emotions Ratings for the Neutral Audio Sequences



Note. Frequencies are normalised between 0 and 1 to represent relative frequencies.

Supplementary Figure 7.10

Frequencies of the GEW Emotions Ratings for the Positive Audio Sequences



Note. Frequencies are normalised between 0 and 1 to represent relative frequencies.

Supplementary Table 7.9

Bootstrapped Average Negative and Positive Basic Emotions per Audio Sequence

Negative Audio Sequences				Neutral Audio Sequences				Positive Audio Sequences			
				Negative Basic Emotions							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	0.77	0.92	1.00	1	0.00	0.08	0.23	1	0.00	0.00	0.00
2	1.00	1.00	1.00	2	0.08	0.31	0.54	2	0.00	0.15	0.38
3	1.00	1.00	1.00	3	0.00	0.15	0.38	3	0.00	0.08	0.23
4	1.00	1.00	1.00	4	0.00	0.15	0.38	4	0.00	0.08	0.23
5	0.77	0.92	1.00	5	0.00	0.15	0.38	5	0.00	0.16	0.38
6	0.77	0.92	1.00	6	0.00	0.00	0.00	6	0.00	0.00	0.00
7	1.00	1.00	1.00	7	0.00	0.15	0.38	7	0.00	0.15	0.38
8	1.00	1.00	1.00	8	0.00	0.08	0.23	8	0.00	0.15	0.38
9	1.00	1.00	1.00	9	0.00	0.08	0.23	9	0.00	0.23	0.46
10	0.77	0.92	1.00	10	0.00	0.08	0.23	10	0.00	0.08	0.23
11	0.77	0.92	1.00	11	0.00	0.23	0.46	11	0.00	0.08	0.23
				12	0.00	0.23	0.46				
				13	0.00	0.15	0.38				
				14	0.00	0.15	0.38				
				15	0.00	0.08	0.23				
Average	0.90	0.97	1.00		0.01	0.14	0.33		0.00	0.10	0.27

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. Positive emotions: joy and surprise (maximum value = 2). Negative emotions: sadness, fear, disgust, and anger (maximum value = 4).

Continued on next page.

Supplementary Table 7.9 ***(continued)***

Bootstrapped Average Negative and Positive Basic Emotions per Audio Sequence

Negative Audio Sequences				**Neutral Audio Sequences**				**Positive Audio Sequences**			
				Positive Basic Emotions							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	0.00	0.08	0.23	1	0.77	0.92	1.00	1	1.00	1.00	1.00
2	0.00	0.00	0.00	2	0.46	0.69	0.92	2	0.62	0.85	1.00
3	0.00	0.00	0.00	3	0.62	0.85	1.00	3	0.77	0.92	1.00
4	0.00	0.00	0.00	4	0.62	0.85	1.00	4	0.77	0.92	1.00
5	0.00	0.08	0.23	5	0.62	0.85	1.00	5	0.62	0.84	1.00
6	0.00	0.08	0.23	6	1.00	1.00	1.00	6	1.00	1.00	1.00
7	0.00	0.00	0.00	7	0.62	0.85	1.00	7	0.62	0.85	1.00
8	0.00	0.00	0.00	8	0.77	0.92	1.00	8	0.62	0.85	1.00
9	0.00	0.00	0.00	9	0.77	0.92	1.00	9	0.54	0.77	1.00
10	0.00	0.08	0.23	10	0.77	0.92	1.00	10	0.77	0.92	1.00
11	0.00	0.08	0.23	11	0.54	0.77	1.00	11	0.77	0.92	1.00
				12	0.54	0.77	1.00				
				13	0.62	0.85	1.00				
				14	0.62	0.85	1.00				
				15	0.77	0.92	1.00				
Average	0.00	0.03	0.10		0.67	0.86	0.99		0.73	0.90	1.00

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. Positive emotions: joy and surprise (maximum value = 2). Negative emotions: sadness, fear, disgust, and anger (maximum value = 4).

Supplementary Table 7.10

Bootstrapped Average Negative and Positive GEW Emotions per Audio Sequence

Negative Audio Sequences				Neutral Audio Sequences				Positive Audio Sequences			
				Negative GEW Emotions							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	2.69	3.54	4.31	1	0.00	0.16	0.46	1	0.00	0.00	0.00
2	3.23	3.92	4.62	2	0.23	1.08	2.15	2	0.00	0.46	1.31
3	3.31	4.16	5.08	3	0.00	0.62	1.54	3	0.00	0.38	1.15
4	4.08	4.61	5.08	4	0.00	0.38	0.92	4	0.00	0.08	0.23
5	2.85	3.70	4.46	5	0.00	0.31	0.77	5	0.00	0.46	1.23
6	3.15	4.00	4.77	6	0.00	0.00	0.00	6	0.00	0.00	0.00
7	3.38	4.07	4.77	7	0.00	0.53	1.31	7	0.00	0.54	1.31
8	3.38	3.93	4.38	8	0.00	0.15	0.46	8	0.00	0.53	1.31
9	3.85	4.46	5.08	9	0.00	0.08	0.23	9	0.00	0.61	1.38
10	2.85	3.76	4.62	10	0.00	0.16	0.46	10	0.00	0.23	0.69
11	3.15	4.08	4.85	11	0.00	0.39	0.85	11	0.00	0.30	0.92
				12	0.00	0.69	1.46				
				13	0.00	0.54	1.46				
				14	0.00	0.38	0.92				
				15	0.00	0.16	0.46				
Average	3.27	4.02	4.73		0.02	0.37	0.90		0.00	0.33	0.87

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. GEW: Geneva Emotion Wheel. Positive emotions: interest, joy, pleasure, satisfaction, amusement, compassion, pride, and relief (maximum value = 8). Negative emotions: fear, disgust, hatred, anger, disappointment, sadness, contempt, shame, guilt, and regret (maximum value = 10). *Continued on next page.*

Supplementary Table 7.10 ***(continued)***

Bootstrapped Average Negative and Positive GEW Emotions per Audio Sequence

Negative Audio Sequences				**Neutral Audio Sequences**				**Positive Audio Sequences**			
				Positive GEW Emotions							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	0.00	0.23	0.69	1	1.31	2.15	3.15	1	4.31	4.77	5.23
2	0.00	0.00	0.00	2	0.69	1.54	2.46	2	2.54	3.62	4.54
3	0.00	0.00	0.00	3	1.46	2.23	3.00	3	3.08	4.08	4.92
4	0.00	0.00	0.00	4	1.23	1.93	2.69	4	3.54	4.53	5.31
5	0.00	0.31	0.92	5	1.69	2.69	3.69	5	2.23	3.31	4.31
6	0.00	0.31	0.92	6	1.92	2.84	3.85	6	3.00	3.92	4.85
7	0.00	0.00	0.00	7	0.85	1.47	2.23	7	1.69	2.54	3.38
8	0.00	0.00	0.00	8	1.62	2.47	3.31	8	2.15	3.16	4.08
9	0.00	0.00	0.00	9	1.31	2.08	2.92	9	2.15	3.24	4.23
10	0.00	0.31	0.92	10	1.62	2.38	3.15	10	2.46	3.55	4.54
11	0.00	0.30	0.92	11	1.00	1.92	2.85	11	2.69	3.62	4.46
				12	0.69	1.23	1.77				
				13	1.23	2.00	2.85				
				14	1.46	2.32	3.15				
				15	1.46	2.23	3.08				
Average	0.00	0.13	0.40		1.30	2.10	2.94		2.71	3.67	4.53

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. GEW: Geneva Emotion Wheel. Positive emotions: interest, joy, pleasure, satisfaction, amusement, compassion, pride, and relief (maximum value = 8). Negative emotions: fear, disgust, hatred, anger, disappointment, sadness, contempt, shame, guilt, and regret (maximum value = 10).

Supplementary Table 7.11

Bootstrapped GEW Emotion Intensity Ratings per Audio Sequence

Negative Audio Sequences				**Neutral Audio Sequences**				**Positive Audio Sequences**			
				GEW Emotion Intensity / Strength							
ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*	ID	*2.5th CI*	*M*	*97.5th CI*
1	3.23	3.77	4.38	1	1.54	2.31	3.23	1	4.31	4.77	5.23
2	3.23	3.92	4.62	2	1.77	2.62	3.46	2	3.31	4.08	4.77
3	3.31	4.16	5.08	3	2.15	2.85	3.54	3	3.69	4.46	5.15
4	4.08	4.61	5.08	4	1.77	2.31	2.92	4	3.77	4.61	5.38
5	3.38	4.00	4.54	5	2.23	3.00	3.85	5	2.92	3.77	4.54
6	3.77	4.31	4.85	6	1.92	2.84	3.85	6	3.00	3.92	4.85
7	3.38	4.07	4.77	7	1.31	2.00	2.77	7	2.38	3.08	3.77
8	3.38	3.93	4.38	8	1.92	2.62	3.38	8	3.08	3.69	4.38
9	3.85	4.46	5.08	9	1.38	2.15	3.00	9	3.15	3.85	4.54
10	3.38	4.07	4.69	10	1.92	2.54	3.15	10	2.92	3.78	4.62
11	3.85	4.39	4.92	11	1.54	2.31	3.08	11	3.15	3.92	4.62
				12	1.38	1.92	2.46				
				13	1.77	2.54	3.38				
				14	2.00	2.70	3.38				
				15	1.77	2.39	3.15				
Average	3.53	4.15	4.76		1.76	2.47	3.24		3.24	3.99	4.71

Note. M: Mean; *CI*: Confidence Interval of the 2.5th and 97.5th percentile. GEW: Geneva Emotion Wheel. Intensity range GEW: 0 – 6.

Supplementary Table 7.12

Results of the Repeated Measure Analyses of Variance (rmANOVA) With Condition (Negative vs. Neutral vs. Positive) as Within-Subject Factor

Rating Type	Dependent Variable	*df1*	*df2*	*F*	*p*	η_p^2
SAM (continuous)	Valence	2	50.0	160.22	$<$.001	0.87
	negative – neutral	-30.51 [-35; -26.07]				
	positive – neutral	19.15 [10.74; 26.35]				
	negative – positive	-49.66 [-56.17; -42.79]				
SAM (continuous)	Arousal	2	50.0	71.73	$<$.001	0.74
	negative – neutral	43.99 [35.27; 52.95]				
	positive – neutral	27.21 [16.98; 37.08]				
	negative – positive	16.78 [10.66; 24.11]				
SAM (post-presentation)	Dominance	2	50	134.41	$<$.001	0.84
	negative – neutral	55.34 [46.31; 63.48]				
	positive – neutral	25.73 [16.55; 34.59]				
	negative – positive	29.61 [24.6; 34.45]				
SAM (continuous)	Valence Standard Deviation	2	50	1.88	.163	0.07
	negative – neutral	1.46 [-0.73; 3.53]				
	positive – neutral	0.56 [-1.01; 2.09]				
	negative – positive	0.9 [-0.61; 2.4]				
SAM (continuous)	Arousal Standard Deviation	2	50	15.20	$<$.001	0.38
	negative – neutral	4.58 [2.24; 7.1]				
	positive – neutral	0.72 [-0.91; 2.42]				
	negative – positive	3.85 [2.01; 5.8]				

Note. Bonferroni-corrected 2.5th and 97.5th confidence intervals (CIs) of the mean obtained via non-parametric bootstrapping with 5000 iterations were used for the post-hoc single comparisons. Contrasts, where the CI does not include zero, can be considered as significant (Cumming & Finch, 2005). *Continued on next page.*

Supplementary Table 7.12 ***(continued)***

Results of the Repeated Measure Analyses of Variance (rmANOVA) With Condition (Negative vs. Neutral vs. Positive) as Within-Subject Factor

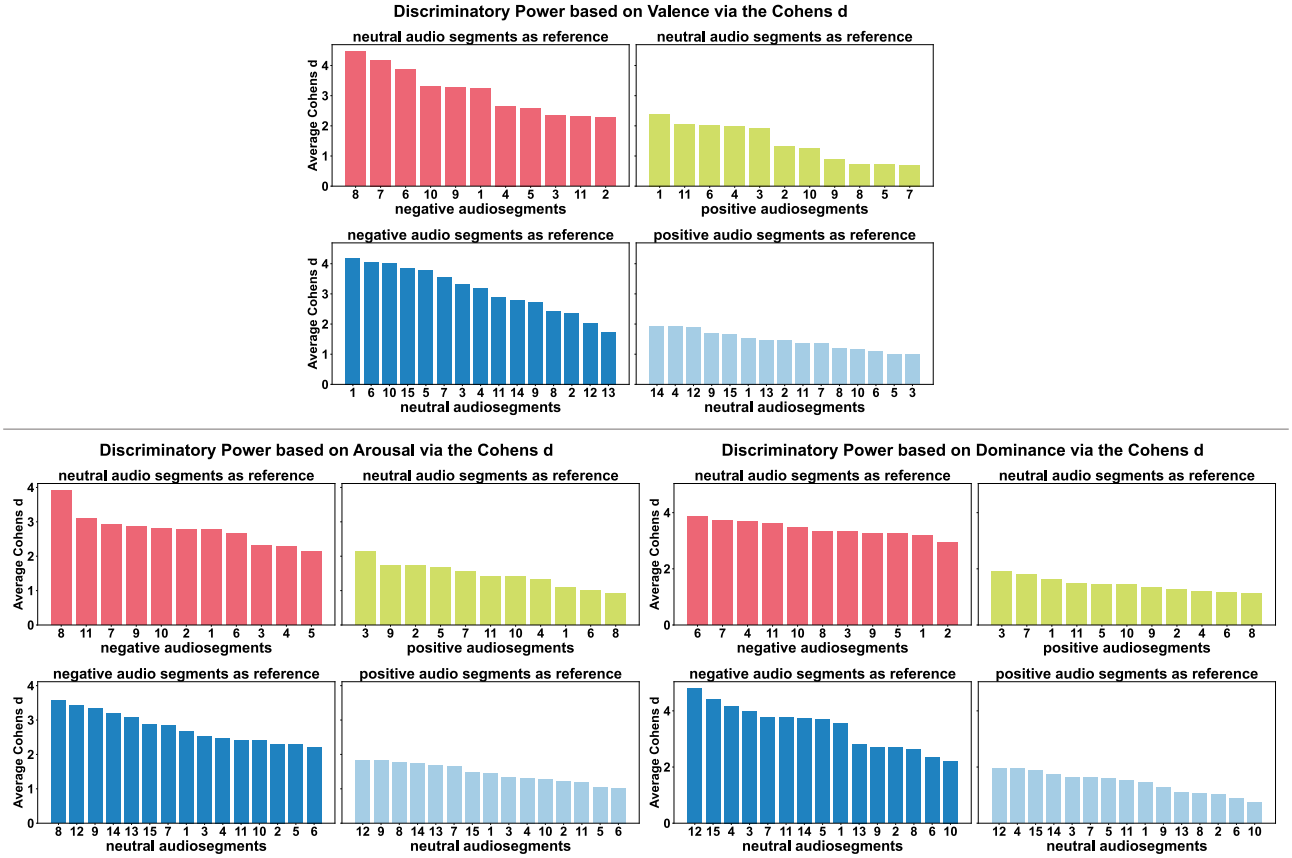
Rating Type	Dependent Variable	*df1*	*df2*	*F*	*p*	η_p^2
Post-presentation rating	Familiarity	2	50	13.89	<.001	0.36
	negative – neutral	-28.87 [-39.62; -18.2]				
	positive – neutral	-7.85 [-23.56; 8.09]				
	negative – positive	-21.01 [-33.94; -8.38]				
Post-presentation rating	Negative Basic Emotions	2	50	170.50	<.001	0.87
	negative – neutral	0.83 [0.71; 0.93]				
	positive – neutral	-0.03 [-0.16; 0.12]				
	negative – positive	0.86 [0.73; 0.95]				
Post-presentation rating	Positive Basic Emotions	2	50	170.50	<.001	0.87
	negative – neutral	-0.83 [-0.93; -0.71]				
	positive – neutral	0.03 [-0.12; 0.16]				
	negative – positive	-0.86 [-0.95; -0.73]				
Post-presentation rating	Negative GEW Emotions	2	50	231.70	<.001	0.90
	negative – neutral	3.64 [3.23; 4.09]				
	positive – neutral	-0.04 [-0.49; 0.52]				
	negative – positive	3.68 [3.23; 4.1]				
Post-presentation rating	Positive GEW Emotions	2	50	70.21	<.001	0.74
	negative – neutral	-1.95 [-2.52; -1.39]				
	positive – neutral	1.6 [0.66; 2.38]				
	negative – positive	-3.55 [-4.11; -2.9]				
Post-presentation rating	GEW Intensity/Strength	2	50	38.60	<.001	0.61
	negative – neutral	1.69 [1.18; 2.18]				
	positive – neutral	1.56 [0.98; 2.12]				
	negative – positive	0.13 [-0.26; 0.54]				

Note. Bonferroni-corrected 2.5^{th} and 97.5^{th} confidence intervals (CIs) of the mean obtained via non-parametric bootstrapping with 5000 iterations were used for the post-hoc single comparisons. Contrasts, where the CI does not include zero, can be considered as significant (Cumming & Finch, 2005).

7.9 Alternative Analysis of the Cohen's D as Measure of Discriminatory Power

Supplementary Figure 7.11

Average Cohen's d as Measure of Discriminatory Power Relative to the Neutral Condition



Note. Average Cohen's d of each emotional audio sequence to the neutral reference condition for the valence (top), arousal (lower, left), and dominance (lower, right) rating.

Averaged or summed effect sizes per audio sequence can be considered as an alternative approach to the here suggested subtractive comparisons via Monte Carlo Simulations (MCS; Ernst, 2004). The effect size Cohen's d was calculated as a measure of emotional distinctiveness between each audio sequence using the following formula:

$$Cohen's\ d = \frac{\text{mean}(x) - \text{mean}(y)}{\sqrt{\frac{\text{std}(x)^2 + \text{std}(y)^2}{2}}} \tag{7.1}$$

The Cohen's d of one audio sequence to each reference condition was averaged, yielding a measure of discriminatory power to the references. Hence, audio sequences with lower effect size are less distinguishable

from the reference condition than those with larger effect size. In general, we observed larger discriminatory power between negative and neutral audio sequences compared to positive and neutral audio sequences. In line with the MCS approach, we observed less discriminatory power regarding valence for the audio sequences positive 7 and positive 8 compared to the neutral reference condition. However, the effect size approach additionally identifies the audio sequence positive 5 as less differentiable. From the neutral condition, the audio sequences neutral 3, neutral 5, and neutral 8 revealed the lowest effect size. Results of the arousal ratings showed low discriminatory power for the audio sequences neutral 6, neutral 5, as well as positive 8. Discriminatory power analysis of the dominance ratings suggested the audio sequences neutral 10, neutral 6, positive 8, and positive 6 as less strong in distinguishing the conditions.

Supplementary References Study II

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *The American Psychologist*, *60*(2), 170–180. https://doi.org/10.1037/0003-066X.60.2.170

Ernst, M. D. (2004). Permutation methods: A basis for exact inference. *Statistical Science*, *19*(4), 676–685. https://doi.org/10.1214/088342304000000396

Janke, S., & Glöckner-Rist, A. (2012). *Deutsche Version der Satisfaction with Life Scale* (SWLS; Manual). ZIS – GESIS Leibniz Institute for the Social Sciences. https://doi.org/10.6102/zis147

Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981). *Das State–Trait–Angstinventar (STAI): Theoretische Grundlagen und Handanweisung* (Manual). Beltz Test GmbH. Weinheim, Germany.

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica*, *51*(4), 195–206. https://doi.org/10.1026/0012-1924.51.4.195

Whooley, M. A., Avins, A. L., Miranda, J., & Browner, W. S. (1997). Case-finding instruments for depression. Two questions are as good as many. *Journal of General Internal Medicine*, *12*(7), 439–445. https://doi.org/10.1046/j.1525-1497.1997.00076.x

World Health Organization. (1998). *World Health Organization Five Well-Being Index (WHO-5)* (Manual; retrieved April 10, 2023). Psychiatric Research Unit, WHO Collaborating Centre in Mental Health, and Frederiksborg General Hospital. Hillerød, Denmark. https://doi.org/10.1037/t36093-000

8 Supplementary Material for Study III

Oscillatory and Gaze Signatures of Emotional Speech Processing, Visuo-Spatial Cognition, and Their Interaction in a Near-Realistic Dual-Task MEG Study

Katharina Lingelbach[1,2], Christoph S. Herrmann[1], and Jochem W. Rieger[1]

[1] Department of Psychology, University of Oldenburg, Oldenburg, Germany

[2] Fraunhofer Institute for Industrial Engineering IAO, Nobelstraße 12, 70569 Stuttgart, Germany

Supplementary Material to the article submitted as:

Lingelbach, K., Herrmann, C. S., & Rieger, J. W. (2025). Oscillatory and gaze signatures of emotional speech processing, visuo-spatial cognition, and their interaction in a near-realistic dual-task MEG study. *Imaging Neuroscience*.

8.1 Stimulus Set From the German AUDItory Emotional Database

We excluded four audio sequences of the neutral condition (i.e., items 3, 5, 8, and 10) from the database to equalise the number of stimuli per condition. Those stimuli were chosen due to their larger confidence interval in their valence ratings in the validation study. The stimuli were previously rated in valence and arousal using a scale between 0 to 100 (Lingelbach et al., 2024). The final negative stimulus set in this study exhibited an average valence of 21.44 ($SD = 10.72$, $min = 2.10$, $max = 53.06$), and an average arousal of 74.21 ($SD = 13.89$, $min = 25.03$, $max = 100.00$). The average valence of neutral stimuli was 50.17 ($SD = 11.29$, $min = 5.46$, $max = 82.53$), while the average arousal was 29.18 ($SD = 18.79$, $min = 0.00$, $max = 77.17$). For positive stimuli, the average valence was 71.15 ($SD = 17.29$, $min = 9.76$, $max = 100.00$), and the average arousal was 57.87 ($SD = 20.95$, $min = 3.17$, $max = 100.00$).

In this study, the valence and arousal ratings of the speech stimuli significantly differed between emotional speech conditions (speech valence: $F_{(2,\ 224.43)} = 89.72$, $p < .001$; excluded outlier observations: $5.49\ \%$; speech arousal: $F_{(2,\ 222.37)} = 22.87$, $p < .001$; excluded outlier observations: $5.49\ \%$; see also Supplementary Figure 8.2C – D), but not between the workload levels. Participants rated positive stimuli significantly higher in valence than neutral (HV – NV: $M = 0.29$, $95\ \%\ CI\ [0.23,\ 0.35]$; $Z_{(47)} = -5.67$, $p < .001$), and negative stimuli (LV – HV: $M = -0.30$, $95\ \%\ CI\ [-0.37,\ -0.22]$; $Z_{(47)} = -4.99$, $p < .001$). Arousal ratings were significantly higher for negative stimuli compared to neutral (LV – NV: $M = 0.16$, $95\ \%\ CI\ [0.09,\ 0.23]$; $Z_{(47)} = -3.71$, $p < .001$), and positive stimuli (LV – HV: $M = 0.17$, $95\ \%\ CI\ [0.11,\ 0.24]$; $Z_{(47)} = -4.53$, $p < .001$). No significant main effects of visuo-spatial workload or further significant pairwise contrasts of emotional speech (valence LV – NV and arousal HV – NV) were observed.

8.2 Analysis of Conversation Content Recall and Misses

Linear mixed-effects models revealed a significant main effect of emotional speech on recall ($F_{(2,\ 225.95)} = 37.18$, $p < .001$; excluded outlier observations: $4.73\ \%$) and content misses ($F_{(2,\ 223.49)} = 12.23$, $p < .001$; excluded outlier observations: $3.60\ \%$), in responses to the conversation content questions answered at the end of each driving block (Supplementary Figure 8.2E – F). Content misses were significantly more frequent for neutral compared to negative conversations (LV – NV: $M = -0.09$, $95\ \%\ CI\ [-0.14,\ -0.04]$; $Z_{(47)} = -3.15$, $p = .005$) and positive conversations (HV – NV: $M = -0.11$, $95\ \%\ CI\ [-0.18,\ -0.04]$; $Z_{(47)} = -2.75$, $p = .008$). Conversely, recall performance was significantly lower for negative compared to neutral conversations (LV – NV: $M = -0.19$, $95\ \%\ CI\ [-0.23,\ -0.14]$; $Z_{(47)} = -5.25$, $p < .001$), and for negative compared to positive conversations (LV – HV: $M = -0.23$, $95\ \%\ CI\ [-0.27,\ -0.18]$; $Z_{(47)} = -5.53$, $p < .001$).

8.3 Centre Frequencies of the Frequency Bands

Supplementary Table 8.1

Descriptive Statistics of Centre Frequencies per Power Band in Hz

Frequency Band	Mean	Standard Deviation	Minimum	Maximum
Alpha	10.29	0.51	9.22	11.46
Beta	19.80	0.25	19.23	20.52
Gamma	33.59	0.78	32.29	35.25

Note. Centre frequencies were derived from the resting state sensor power through power parametrisation using the FOOOF algorithm (Donoghue et al., 2020).

8.4 Common Spatial Patterns for Multiclass Decoding

Common spatial pattern (CSP) analysis provides a data-driven, supervised signal decomposition that extracts components maximally discriminative in signal variance between experimental conditions. It computes spatial filters that maximise variance in one class (i.e., experimental condition), while reducing it in others (Blankertz et al., 2008). Its extension to multiclass decoding is achieved via joint approximate diagonalisation (JAD) of class-wise covariance matrices (Grosse-Wentrup & Buss, 2008). JAD estimates a transformation matrix that approximately diagonalises these covariance matrices across all classes. The columns of the obtained transformation matrix serve as spatial filters and are ranked to maximise the mutual information between the extracted components and class labels. Each column of the transformation matrix defines a spatial filter. Applying these filters to the data yields components, which are subsequently ranked by their mutual information with the class labels (Grosse-Wentrup & Buss, 2008). These CSP components serve as features for subsequent machine learning classification, for example, using Linear Discriminant Analysis (LDA), to distinguish between the different classes. Similar to the CSP, LDA projects the input data onto a linear subspace that maximises separation between classes.

8.5 Calculation of Gaze Dispersion, Pupil Diameter in Millimetres, and Pupillary Activity Index

8.5.1 Gaze Dispersion

To quantify gaze dispersion, we first computed the gaze centroid (x_c, y_c) as the x- and y-coordinates averaged across all fixations (x_i, y_i: horizontal and vertical coordinates of fixation i) within the block. Fixations were weighted by their duration (d_i: duration of fixation i).

$$x_c = \frac{\sum_i d_i x_i}{\sum_i d_i}, \quad y_c = \frac{\sum_i d_i y_i}{\sum_i d_i} \tag{8.1}$$

Next, we calculated the Euclidean distance between each fixation i and the centroid (x_c, y_c).

$$D_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \tag{8.2}$$

In the last step, gaze dispersion was calculated as the square root of the duration-weighted average of squared Euclidean distance D_i^2, where D_i is the distance between fixation i and the centroid, and d_i is the duration of fixation i.

$$\text{RMS} = \sqrt{\frac{\sum_i d_i D_i^2}{\sum_i d_i}} \tag{8.3}$$

8.5.2 Conversion of Pupil Diameter to Millimetres

The raw pupil diameter values were recorded in arbitrary units (AU), defined as the number of pixels on the eye-tracking camera corresponding to the pupil size. To convert these to millimetres, we measured an artificial eye with a known, fixed pupil size (8 mm) to obtain the corresponding number of pixels (2494 pixels). This allowed us to compute a scaling factor:

$$\text{Scaling factor} = \frac{8\,\text{mm}}{\sqrt{2494}} = 0.160 \tag{8.4}$$

This factor was then applied to the data using the following transformation:

$$\text{Pupil diameter (mm)} = \text{Scaling factor} \times \sqrt{\text{Pupil diameter (AU)}} \tag{8.5}$$

8.5.3 Index of Pupillary Activity (IPA)

To compute the index of pupillary activity (IPA), we segmented the pupil diameter time series of each driving block into 1-minute epochs with $50\ \%$ overlap, and downsampled the data from 1000 Hz to 250 Hz, applying zero-phase low-pass filtering beforehand. Pupil diameter signals were decomposed using a two-level periodic Discrete Wavelet Transform (DWT; symlet-16 from the Daubechies wavelet family; pywavelets 1.8.0). In the second step, abrupt fluctuations were detected by identifying local maxima in the wavelet modulus of high-frequency coefficients extracted from the second level of decomposition (Duchowski, 2018). To filter out noise, the wavelet modulus maxima coefficients were thresholded using universal thresholding, defined as

$$\lambda_{\text{univ}} \ = \ \hat{\sigma}\sqrt{2\log n}, \tag{8.6}$$

where $\hat{\sigma}$ is the standard deviation of the modulus maxima values and n is the number of data points (Duchowski, 2018). In the final step, the IPA score was computed by normalising the count of remaining peaks exceeding the threshold by the total signal duration $(t_2 - t_1)$:

$$\text{IPA} \ = \ \frac{\text{Number of abrupt fluctuations}}{\text{Signal duration (s)}} \tag{8.7}$$

Intermediate steps involved in the computation of the IPA are illustrated in Supplementary Figure 8.1.

8.6 Detailed Summary of the Linear Mixed-Effects Models

Supplementary Table 8.2

Summary of the Linear Mixed-Effects Models for Subjective Ratings Analysing Main Effects of Emotional Speech and Visuo-Spatial Workload, as Well as Their Interaction

Effects	*df1*	*df2*	*F*	*p*	Post-Hoc Wilcoxon Signed-Rank Test
Valence					
Visuo-spatial Workload	1	207.34	7.19	.008**	
Emotional Speech	2	207.76	11.34	$<$ .001***	
HV – NV					$Z_{(46)} = -4.23, p < .001$***
Emotional Speech $\times$ Visuo-spatial Workload	2	206.49	4.40	.013*	
LV – NV $\times$ High – Low					Model CI $[0.11, 1.11]$
HV – NV $\times$ High – Low					Model CI $[-0.57, 0.39]$
LV – HV $\times$ High – Low					Model CI $[0.20, 1.20]$
Arousal					
Visuo-spatial Workload	1	222.35	7.24	.008**	
Emotional Speech	2	222.36	4.22	.016*	
LV – NV					$Z_{(47)} = -0.37, p = .711$
HV – NV					$Z_{(47)} = -1.94, p = .079$
Emotional Speech $\times$ Visuo-spatial Workload	2	222.00	3.57	.030*	
LV – NV $\times$ High – Low					Model CI $[-0.63,\ 0.23]$
HV – NV $\times$ High – Low					Model CI $[-0.06,\ 0.82]$
LV – HV $\times$ High – Low					Model CI $[-1.00,\ -0.13]$
Frustration					
Visuo-spatial Workload	1	220.55	40.57	$<$ .001***	
Emotional Speech	2	221.28	9.72	$<$ .001***	
HV – NV					$Z_{(47)} = -4.24, p < .001$***
Emotional Speech $\times$ Visuo-spatial Workload	2	220.85	7.15	.001**	
LV – NV $\times$ High – Low					Model CI $[-1.07,\ -0.25]$
HV – NV $\times$ High – Low					Model CI $[-0.39,\ 0.41]$
LV – HV $\times$ High – Low					Model CI $[-1.08,\ -0.27]$

Note. Model confidence intervals (CI) for significant interaction effects were estimated via bootstrapping of the linear mixed-effects model using 5,000 iterations. Directions of the experimental modulations are visualised in Figure 4.2. In the presence of a significant interaction, only interpretable main effects were reported. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload. Significance levels: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

Continued on next page.

Supplementary Table 8.2 ***(continued)***

Summary of the Linear Mixed-Effects Models for Subjective Ratings Analysing Main Effects of Emotional Speech and Visuo-Spatial Workload, as Well as Their Interaction

Effects	*df1*	*df2*	*F*	*p*	Post-Hoc Wilcoxon Signed-Rank Test
Effort					
Visuo-spatial Workload	1	207.24	18.42	$< .001^{***}$	
HW – LW					$Z_{(47)} = -4.48, p < .001^{***}$
Emotional Speech	2	206.97	4.82	.009**	
LV – NV					$Z_{(47)} = -0.16, p = .874$
HV – NV					$Z_{(47)} = -3.50, p = .001^{**}$
LV – HV					$Z_{(47)} = -2.75, p = .009^{**}$
Emotional Speech × Visuo-spatial Workload	2	206.12	1.63	.199	
Distraction					
Visuo-spatial Workload	1	193.77	0.07	.791	
Emotional Speech	2	195.89	5.90	.003**	
LV – NV					$Z_{(44)} = -2.88, p = .012^{*}$
HV – NV					$Z_{(44)} = -1.96, p = .050$
LV – HV					$Z_{(44)} = -2.22, p = .040^{*}$
Emotional Speech × Visuo-spatial Workload	2	194.55	0.49	.613	

Note. Model confidence intervals (CI) for significant interaction effects were estimated via bootstrapping of the linear mixed-effects model using 5,000 iterations. Directions of the experimental modulations are visualised in Supplementary Figure 8.2A – B. Significance levels: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Table 8.3

Summary of the Linear Mixed-Effects Models for Supplementary Subjective Ratings Analysing Main Effects of Emotional Speech and Visuo-Spatial Workload, as Well as Their Interaction

Effects	*df1*	*df2*	*F*	*p*	Post-Hoc Wilcoxon Signed-Rank Test
Speech Valence					
Visuo-spatial Workload	1	222.67	0.00	.970	
Emotional Speech	2	224.43	89.72	$<$ .001***	
LV – NV					$Z_{(47)} = -0.65, p = .525$
HV – NV					$Z_{(47)} = -5.67, p < .001$***
LV – HV					$Z_{(47)} = -4.99, p < .001$***
Emotional Speech × Visuo-spatial Workload	2	222.09	1.28	.281	
Speech Arousal					
Visuo-spatial Workload	1	221.33	0.48	.490	
Emotional Speech	2	222.37	22.87	$<$ .001***	
LV – NV					$Z_{(47)} = -3.71, p < .001$***
HV – NV					$Z_{(47)} = -0.42, p = .672$
LV – HV					$Z_{(47)} = -4.53, p < .001$***
Emotional Speech × Visuo-spatial Workload	2	221.30	0.48	.618	
Content Misses (%)					
Visuo-spatial Workload	1	222.61	0.20	.652	
Emotional Speech	2	223.49	12.23	$<$ .001***	
LV – NV					$Z_{(47)} = -3.15, p = .005$**
HV – NV					$Z_{(47)} = -2.75, p = .008$**
LV – HV					$Z_{(47)} = -1.11, p = .269$
Emotional Speech × Visuo-spatial Workload	2	222.58	0.79	.457	
Recall Accuracy (%)					
Visuo-spatial Workload	1	226.00	0.94	.333	
Emotional Speech	2	225.95	37.18	$<$ .001***	
LV – NV					$Z_{(47)} = -5.25, p < .001$***
HV – NV					$Z_{(47)} = -1.43, p = .154$
LV – HV					$Z_{(47)} = -5.53, p < .001$***
Emotional Speech × Visuo-spatial Workload	2	226.05	0.83	.438	

Note. Model confidence intervals (CI) for significant interaction effects were estimated via bootstrapping of the linear mixed-effects model using 5,000 iterations. Directions of the experimental modulations are visualised in Supplementary Figure 8.2C – F. Significance levels: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Table 8.4

Summary of the Linear Mixed-Effects Models for Driving Quality and Measures of Gaze Behaviour Analysing Main Effects of Emotional Speech and Visuo-Spatial Workload, as Well as Their Interaction

Effects	*df1*	*df2*	*F*	*p*	Post-Hoc Wilcoxon Signed-Rank Test
Driving Quality					
Visuo-spatial Workload	1	215.53	263.22	$< .001^{***}$	
HW – LW					$Z_{(47)} = -6.03, p < .001^{***}$
Emotional Speech	2	205.69	0.46	.633	
Emotional Speech × Visuo-spatial Workload	2	205.78	0.47	.625	
Gaze Dispersion					
Visuo-spatial Workload	1	181.51	9.95	$.002^{**}$	
HW – LW					$Z_{(41)} = -2.82, p = .004^{**}$
Emotional Speech	2	180.53	1.69	.188	
Emotional Speech × Visuo-spatial Workload	2	180.39	1.22	.297	
Blink Activity					
Visuo-spatial Workload	1	179.92	0.37	.543	
Emotional Speech	2	180.12	4.53	$.012^{*}$	
LV – NV					$Z_{(41)} = -0.21, p = .843$
HV – NV					$Z_{(41)} = -1.31, p = .587$
LV – HV					$Z_{(41)} = -0.53, p = .843$
Emotional Speech × Visuo-spatial Workload	2	179.84	0.50	.610	
Pupil Dilation					
Visuo-spatial Workload	1	182.88	37.53	$< .001^{***}$	
HW – LW					$Z_{(41)} = -4.67, p < .001^{***}$
Emotional Speech	2	179.35	3.95	$.021^{*}$	
LV – NV					$Z_{(41)} = -1.21, p = .349$
HV – NV					$Z_{(41)} = -1.44, p = .349$
LV – HV					$Z_{(41)} = -0.32, p = .757$
Emotional Speech × Visuo-spatial Workload	2	180.96	0.16	.854	
Index of Pupillary Activity (IPA)					
Visuo-spatial Workload	1	183.76	0.42	.516	
Emotional Speech	2	184.55	0.74	.479	
Emotional Speech × Visuo-spatial Workload	2	184.78	1.74	.178	

Note. Driving quality is an aggregated score, computed as the average of the root mean square of successive differences in acceleration, steering wheel angle, mean brake actuation, mean deviation from the lane centre, and the number of traffic rule violations (Supplementary Figure 8.3F – J). We observed significant main effects of visuo-spatial workload in all subvariables of the aggregated score (acceleration RMSSD: $F_{(1,\ 205.00)} = 54.93$, $p < .001$; steering wheel angle RMSSD: $F_{(1,\ 210.35)} = 103.26$, $p < .001$; mean brake actuation: $F_{(1,\ 204.12)} = 106.27$, $p < .001$; mean deviation from lane centre: $F_{(1,\ 206.44)} = 25.14$, $p < .001$; number of traffic rule violations: $F_{(1,\ 230.42)} = 276.45$, $p < .001$), but no main effects of emotional speech or interaction effects. This validates their aggregation. Directions of the experimental modulations are visualised in Figure 4.4C – F, and Supplementary Figure 8.3A – E. Significance levels: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

8.7 Aperiodic Component Localised in Source Space With DICS

Modulations in components of the $1/f$-like aperiodic part of the power spectrum have recently gained attention, and their links to cognitive processes have been investigated. For instance, changes in the aperiodic exponent have been proposed to be involved in the synaptic balance between excitation and inhibition, with a more negative (steeper) slope parameter β, indicating increased inhibition (Lu et al., 2024). Additionally, aperiodic broadband power was found to decrease across the brain with increasing cognitive demand (Lu et al., 2024). C. Zhang et al. (2023) proposed a positive relationship between aperiodic components and cognitive control involved in the stable allocation of resources, as opposed to flexible resource allocation. It is further associated with neurophysiological changes across the lifespan (Thuwal et al., 2021), arousal states, and auditory stimulus processing (Jacob et al., 2021).

Based on these prior findings, we explored the effects of emotional speech and visuo-spatial workload on the aperiodic offset, slope, and broadband. An increase in aperiodic components was expected in response to internalised processing and emotional speech, and a decrease during externalised processing and high workload.

The permutation-based clustering revealed a significant main effect of emotional speech in the offset (rh: $p = .034$) and broadband (lh: $p = .024$; rh: $p = .010$; Supplementary Figure 8.8A,B). No significant clusters were found for the aperiodic slope component or the main effect of visuo-spatial workload.

The cluster of the aperiodic offset included vertices in the right temporo-parietal junction (TPJ). The broadband clusters were localised bilaterally in the supramarginal gyrus, TPJ, and extended into the superior and middle temporal sulcus.

Post-hoc tests showed increased aperiodic values during emotional compared to neutral speech (Supplementary Figure 8.8C,D). Offset values were significantly higher for both low (LV – NV: $M = 0.04$, $95\ \%\ CI\ [0.02,\ 0.06]$; $Z_{(41)} = -5.36$, $p < .001$), and high valence speech (HV – NV: $M = 0.04$, $95\ \%\ CI\ [0.02,\ 0.06]$; $Z_{(41)} = -5.14$, $p < .001$). Similarly, broadband values increased for low (LV – NV: $M = 0.05$, $95\ \%\ CI\ [0.03,\ 0.07]$; $Z_{(41)} = -5.30$, $p < .001$), and high valence speech (HV – NV: $M = 0.07$, $95\ \%\ CI\ [0.04,\ 0.10]$; $Z_{(41)} = -5.68$, $p < .001$) compared to neutral speech.

8.8 Modulations in the Non-Corrected Band Source Power Localised With DICS

When the aperiodic component is not separated from the power spectrum, permutation-based clustering reveals a main effect of emotional speech that closely resembles the effect observed in the aperiodic component analysis. Significant differences between emotional speech conditions were found in alpha- (left: $p = .007$; right: $p = .001$), beta- (left: $p < .001$; right: $p < .001$), and gamma-band power (right: $p = .03$). Similar to the patterns observed in the aperiodic offset and broadband effects (Supplementary Analysis 8.7), the clusters included the SMG, TPJ, and the posterior superior temporal sulcus (STS), with stronger effects observed in the right hemisphere. As with the aperiodic offset and broadband effects, post-hoc tests of average

source power across significant clusters showed increased alpha (LV – NV: $M = 0.07, 95\,\% CI\,[0.04,\ 0.11]$; $Z_{(41)} = -5.32, p < .001$; HV – NV: $M = 0.08, 95\,\% CI\,[0.04,\ 0.11]$; $Z_{(41)} = -5.37, p < .001$), and beta power (LV – NV: $M = 0.07, 95\,\% CI\,[0.04,\ 0.10]$; $Z_{(41)} = -5.30, p < .001$; HV – NV: $M = 0.07, 95\,\% CI\,[0.04,\ 0.10]$; $Z_{(41)} = -5.34, p < .001$) during both negative and positive speech compared to neutral speech.

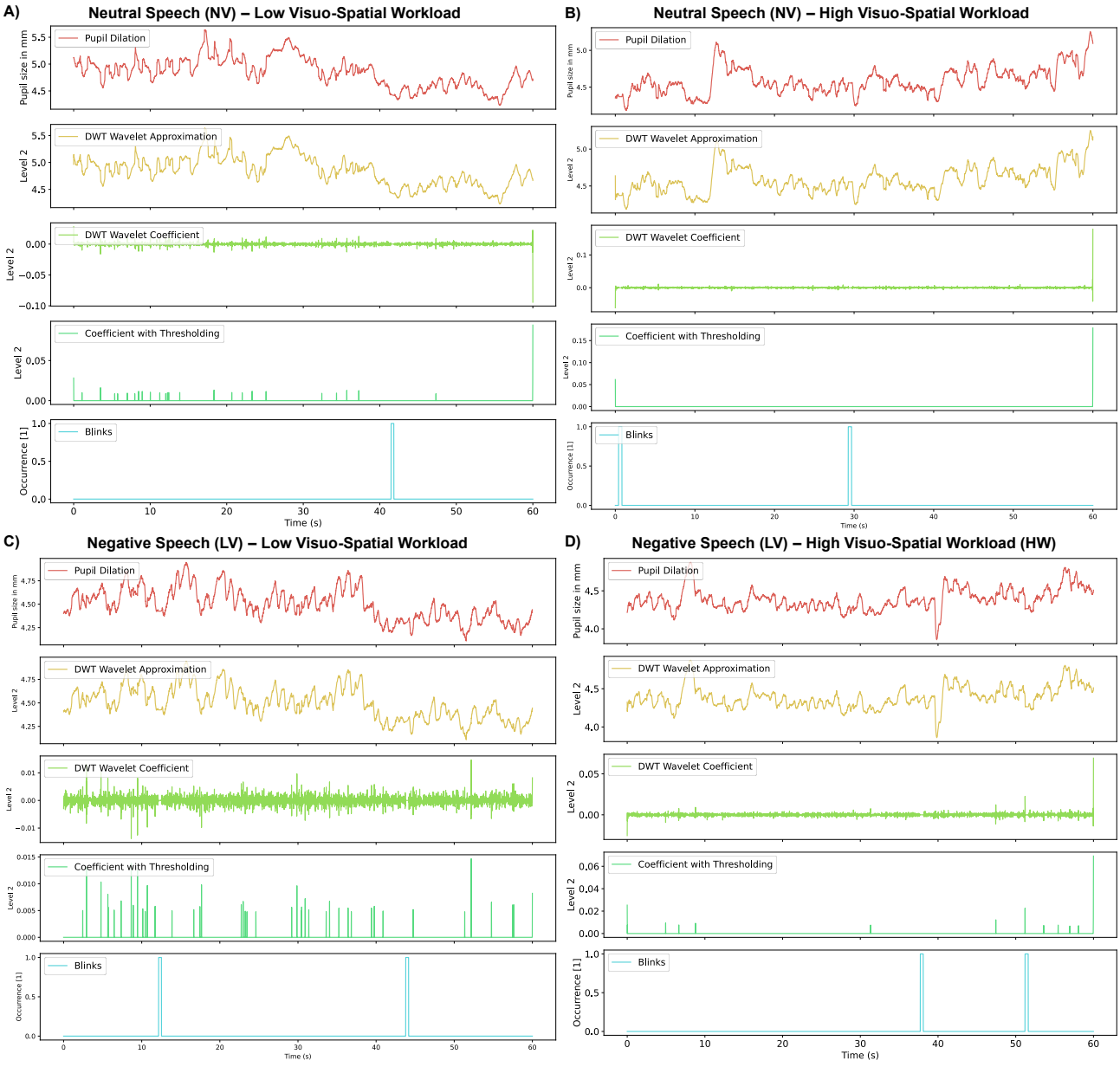
In the gamma band, power in the posterior STS and middle temporal sulcus increased for positive relative to neutral speech (HV – NV: $M = 0.07, 95\,\% CI\,[0.03,\ 0.10]$; $Z_{(41)} = -4.93, p < .001$), and negative relative to positive speech (LV – HV: $M = -0.05, 95\,\% CI\,[-0.09,\ -0.01]$; $Z_{(41)} = -2.67, p = .011$).

The close similarity to the aperiodic effects suggests that, without separating the $1/f$ component, changes in aperiodic activity may obscure oscillatory modulations in the frequency peak regions (above the continuous, aperiodic component).

8.9 Supplementary Figures

Supplementary Figure 8.1

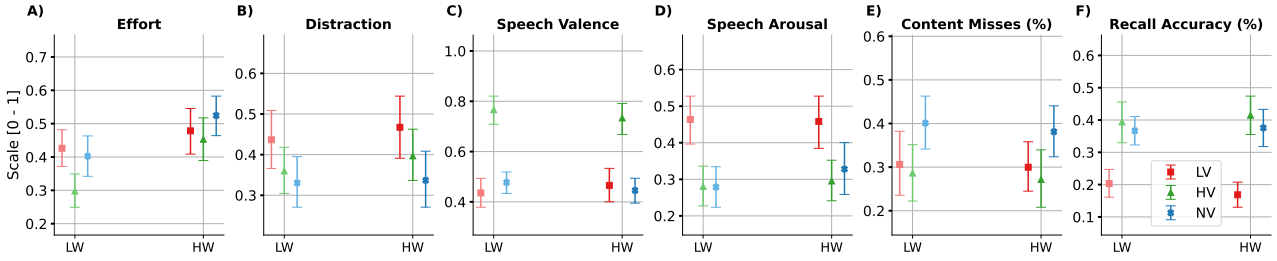
Visualisation of the Intermediate Steps Involved in the Computation of the Index of Pupillary Activity (IPA)



Note. The process is demonstrated using four illustrative conditions. Conditions were selected to represent the most pronounced differences in auditory complexity/arousal and visuo-spatial workload. Pupil dilation (top row) and blink occurrence (bottom row) are visualised alongside the wavelet approximation and coefficients. DWT: Discrete Wavelet Transform.

Supplementary Figure 8.2

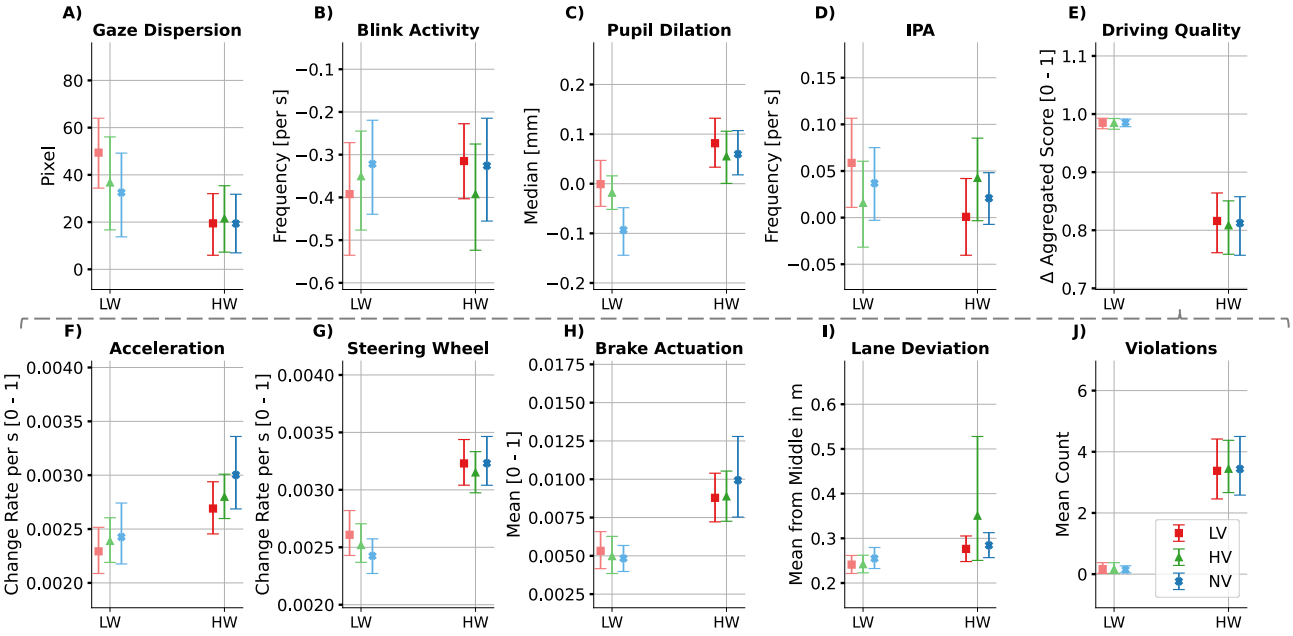
Descriptive Statistics for Subjective Ratings (Excluding Ratings With a Significant Interaction Effect)



Note. Coloured symbols and error bars represent the bootstrapped grand averages for each condition, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI) across participants. LV: low valence; NV: neutral valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.3

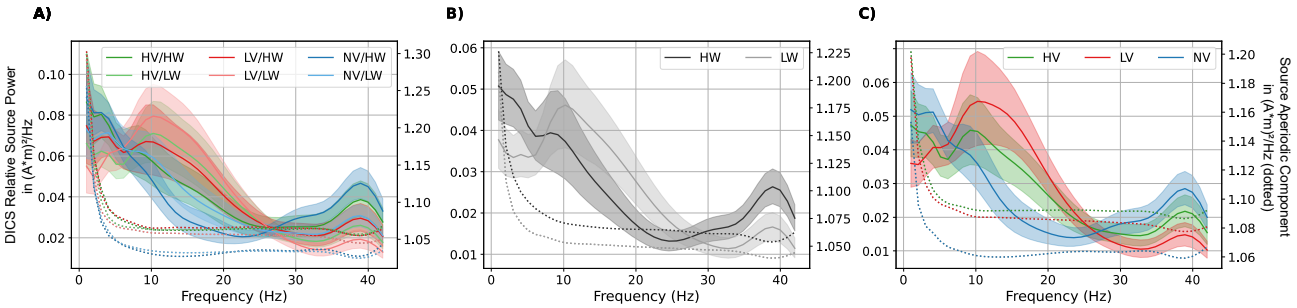
Descriptive Statistics for Correlates of Gaze and Driving Behaviour



Note. Coloured symbols and error bars represent the bootstrapped grand averages for each condition, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI) across participants. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.4

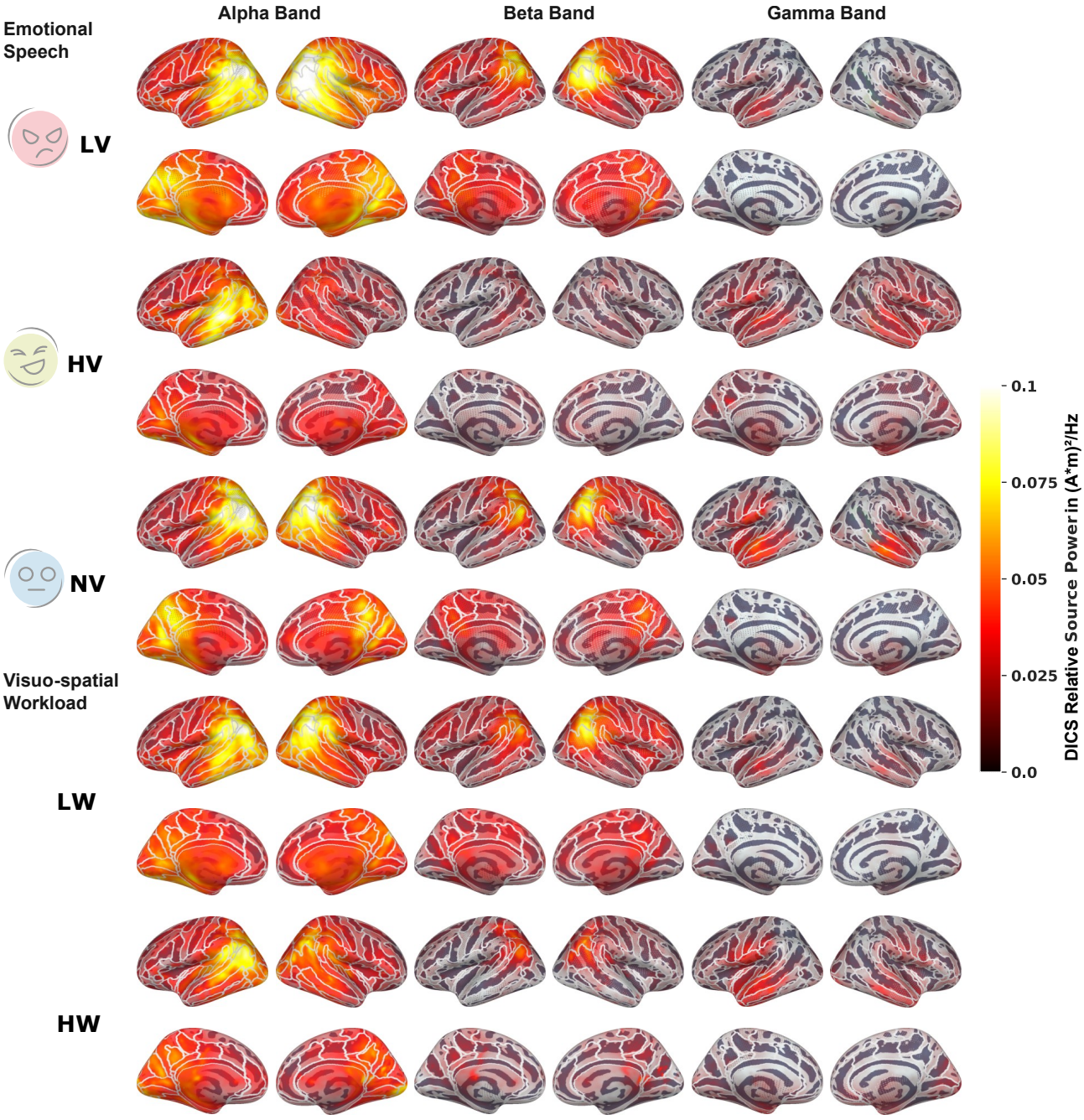
Frequency Spectra of Source-Localised Oscillatory Power Obtained Using Dynamic Imaging of Coherent Sources (DICS)



Note. Panels (A) to (C) show the oscillatory source power when subtracting the aperiodic component (solid lines) and estimated aperiodic component (dotted lines). Spectra correspond to the conditions of (A) single effects, (B) main effect of visuo-spatial workload, and (C) main effects of emotional speech. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.5

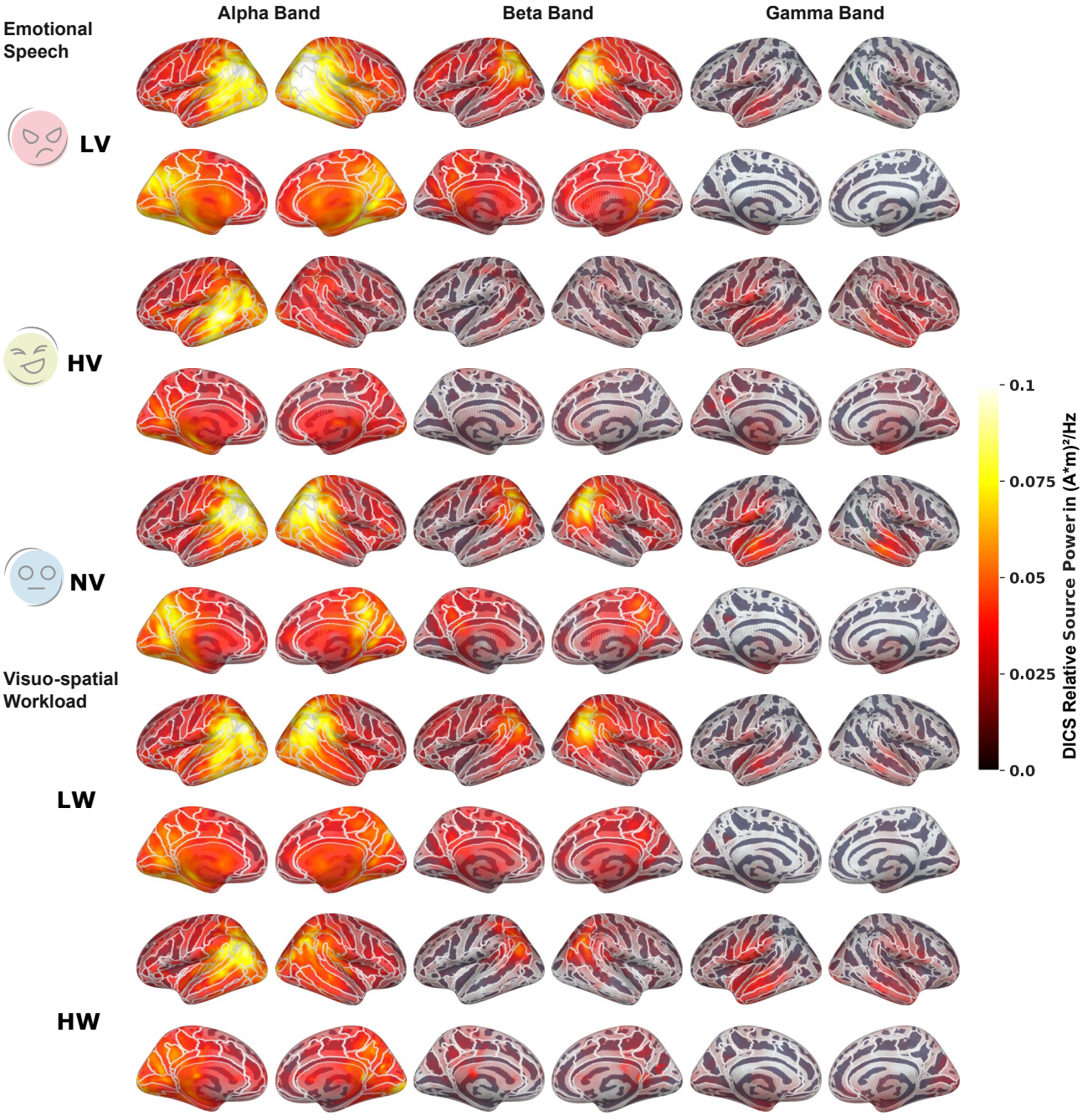
Oscillatory Source Power After Subtraction of the Aperiodic Component per Frequency Band in the Main Effect Conditions



Note. Source power was estimated using dynamic imaging of coherent sources (DICS). LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.6

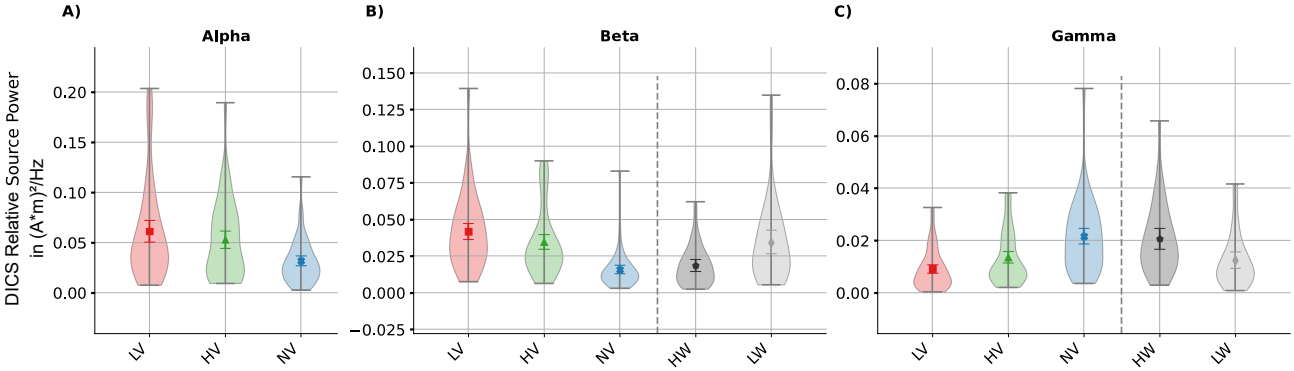
Oscillatory Source Power Without Subtraction of the Aperiodic Component per Frequency Band in the Main Effect Conditions



Note. Source power was estimated using dynamic imaging of coherent sources (DICS). LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.7

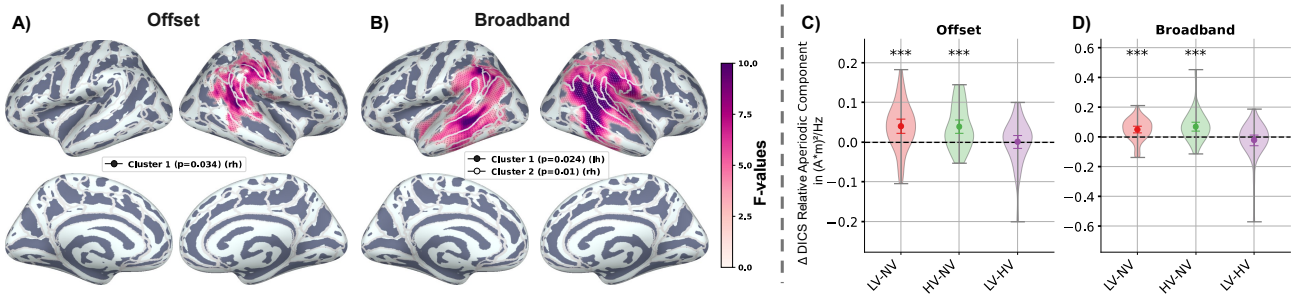
Descriptive Statistics for Source Power in Frequency Bands Within Significant Clusters



Note. Source power was estimated using dynamic imaging of coherent sources (DICS). Coloured dots and error bars represent the bootstrapped grand averages for each condition, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI) across participants. The shaded area indicates the estimated probability density of the data. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.8

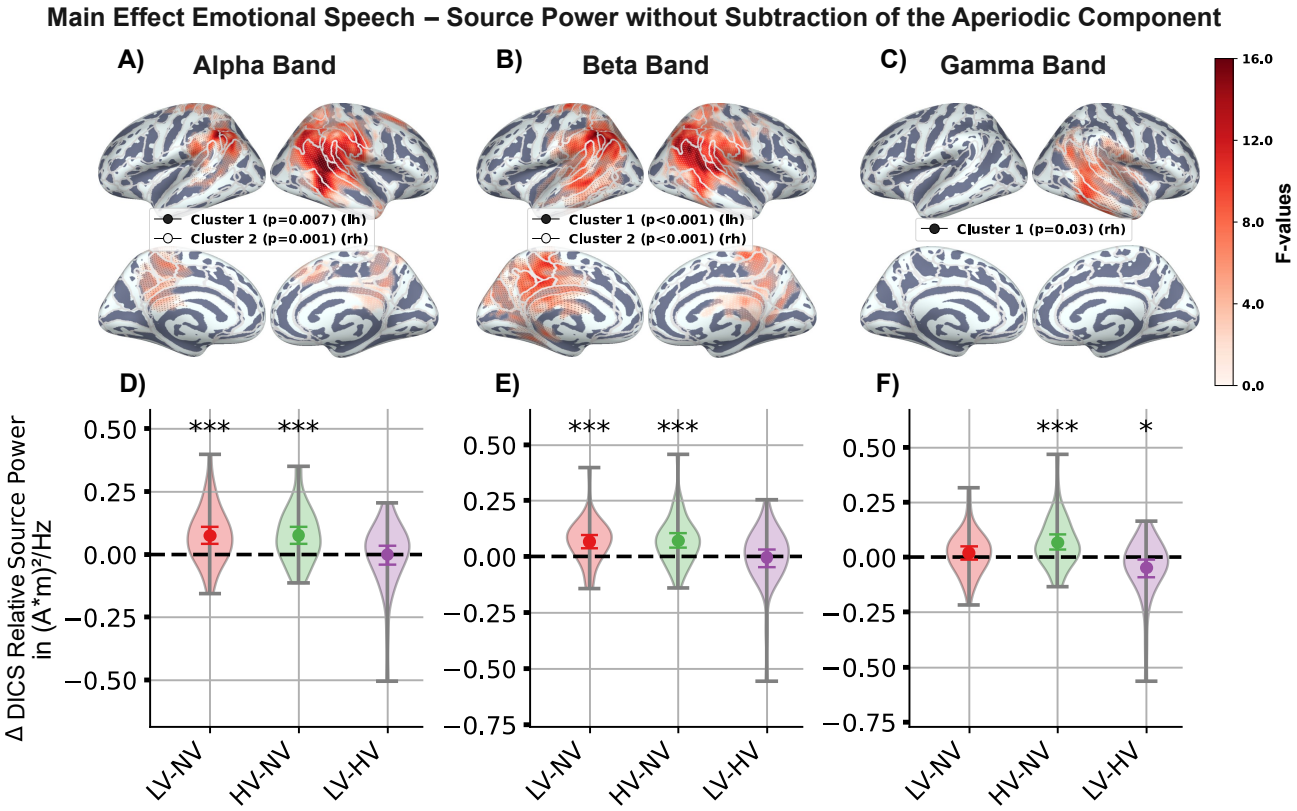
Permutation-Based Spatial Clustering Results of the Main Effect Emotional Speech on the Aperiodic Components



Note. Spatial distribution of significant *F*-values from the *F*-test clusters for the main effect of emotional speech is shown in Panel (A) for the offset and (B) for the broadband. Source power was estimated using dynamic imaging of coherent sources (DICS). Panels (C) and (D) show the contrasts of the post-hoc comparisons conducted using Wilcoxon signed-rank tests (FDR-corrected). Contrasts represent the difference between the conditions in the averaged aperiodic component values from significant vertices. Coloured dots and error bars represent the bootstrapped grand averages for each contrast, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI) across participants. The shaded area indicates the estimated probability density of the data. Significance levels: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.9

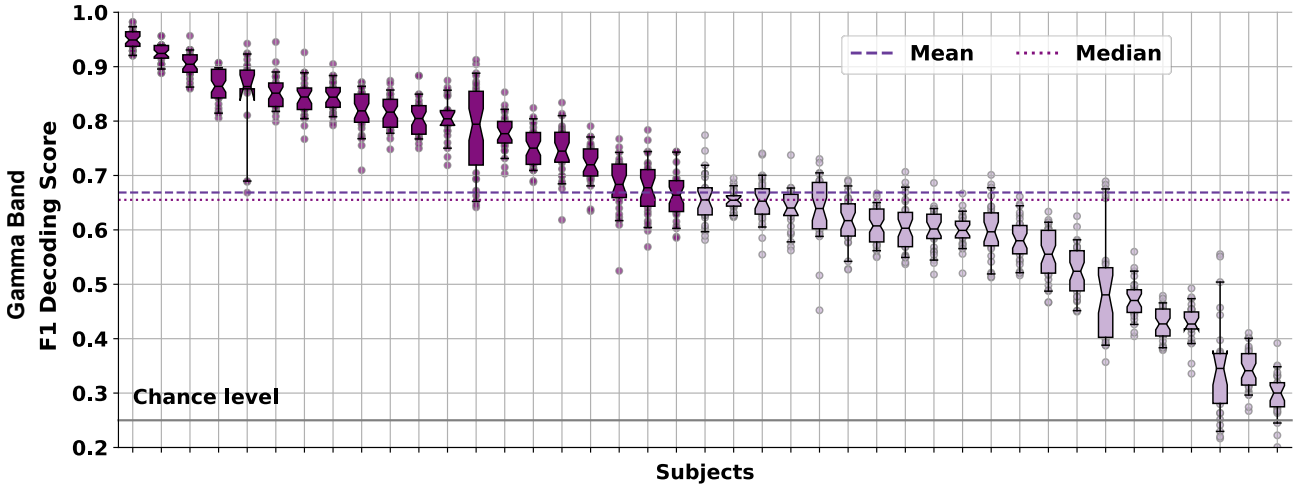
Permutation-Based Spatial Clustering Results of the Main Effect Emotional Speech Performed on the Source Power Without Subtracting the Aperiodic Component



Note. Spatial distribution of significant *F*-values from the *F*-test clusters for the main effect of emotional speech on (A) alpha-, (B) beta-, and (C) gamma-band power. Source power was estimated using dynamic imaging of coherent sources (DICS). Panels (D) to (F) show the contrasts of the post-hoc comparisons conducted using Wilcoxon signed-rank tests (FDR-corrected). Contrasts represent the difference between the conditions in the averaged band power from significant vertices. Coloured dots and error bars represent the bootstrapped grand averages for each contrast, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI) across participants. The shaded area indicates the estimated probability density of the data. Significance levels: * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. LV: low valence; NV: neutral valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.10

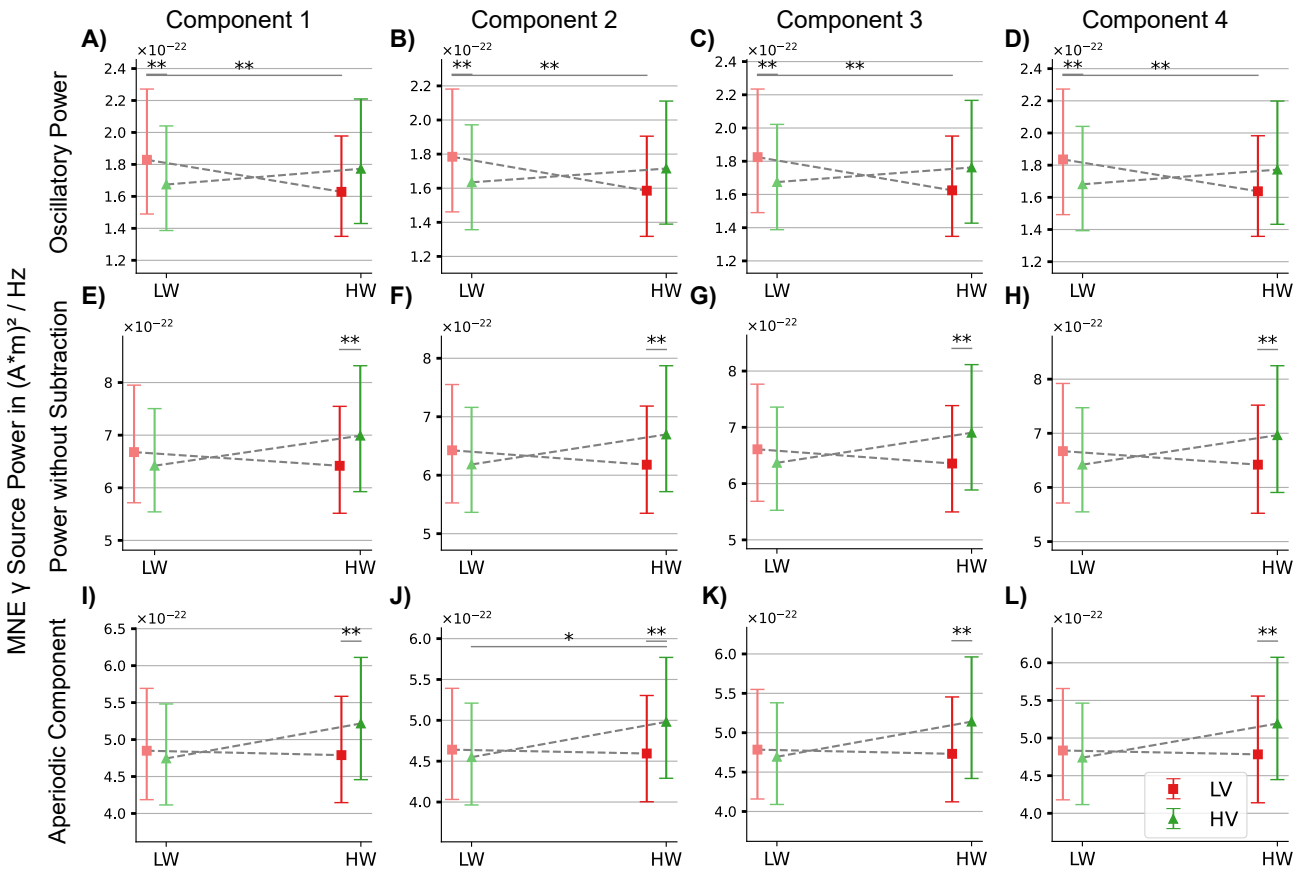
Bootstrapped Mean F1 Scores Across Folds per Subject



Note. Notches in the boxplots indicate the $95\ \%$ confidence interval of the bootstrapped mean. The box spans the interquartile range (25^{th} to 75^{th} percentile), and the whiskers represent the 5^{th} and 95^{th} percentiles. The purple dashed line indicates the overall mean, the purple dotted line the median, and the grey line the chance level at 0.25.

Supplementary Figure 8.11

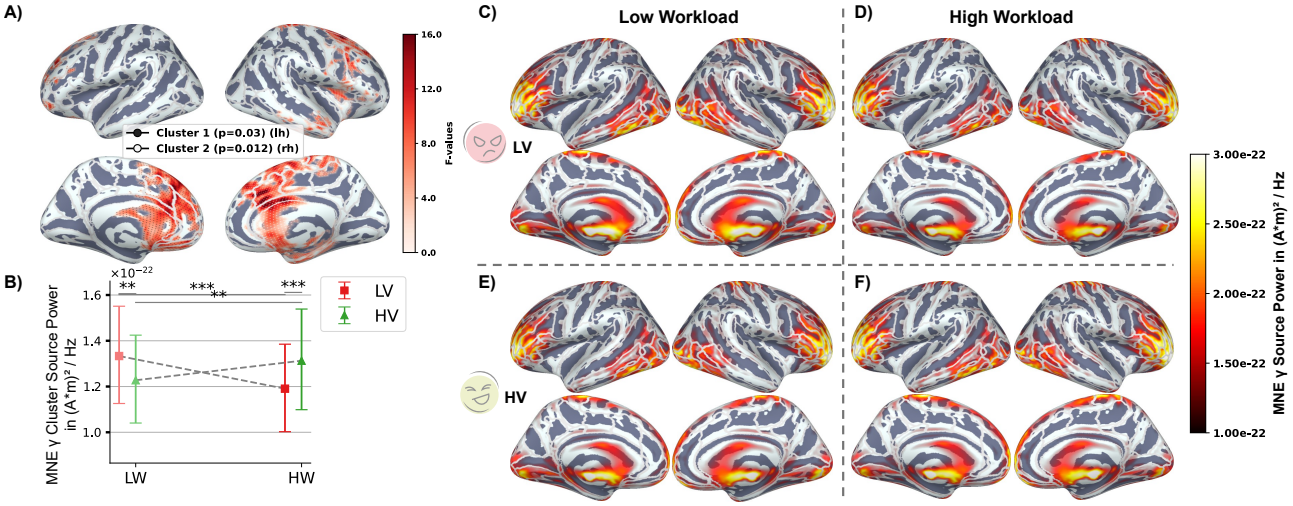
Average Gamma Source Power Across Vertices With Power Exceeding the 90th Percentile From the Decoding Components of the Common Spatial Pattern Algorithm



Note. The CSP components were used in the four-class classification of the interaction. Panels (A) – (D) show the gamma source power localised with MNE and corrected with the aperiodic component. Panels (E) – (H) show the gamma source power without separation of the aperiodic component. Panels (I) – (L) show the aperiodic broadband. Post-hoc comparisons were conducted using Wilcoxon signed-rank tests (FDR-corrected). Coloured symbols and error bars represent the bootstrapped grand averages for each condition, along with their Bonferroni-corrected 2.5th and 97.5th percentile confidence intervals (CI) across participants. Significance levels from the Wilcoxon signed-rank test: *** for $p < .001$, ** for $p < .01$, * for $p < .05$. LV: low valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.12

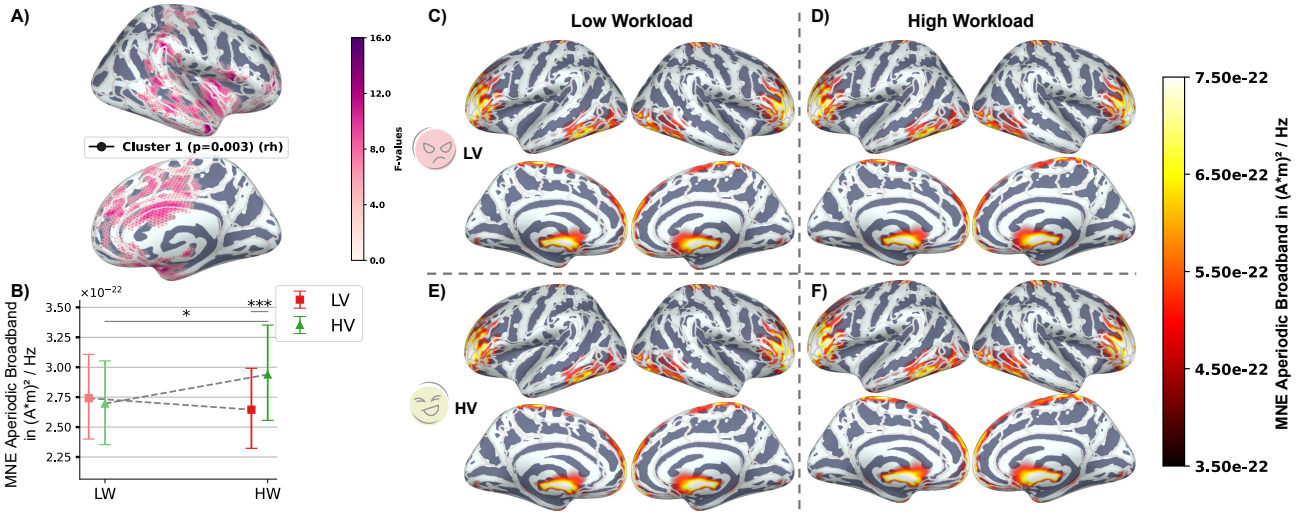
Permutation-Based Spatial Clustering Results of the Gamma-Band Interaction Effect and Source Estimates



Note. Spatial distribution of significant *F*-values of the clustering is visualised in Panel (A). Clustering was performed on oscillatory gamma-band power after subtracting the aperiodic component. Source power was estimated using the minimum norm estimation (MNE) method. Panel (B) shows the gamma-band source power within significant vertices for the conditions. Post-hoc comparisons were conducted using Wilcoxon signed-rank tests (FDR-corrected). Coloured symbols and error bars represent the bootstrapped grand averages for each condition, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI) across participants. Panels (C) to (F) display the source gamma-band power per condition. Significance levels from the Wilcoxon signed-rank test: *** for $p < .001$, ** for $p < .01$, * for $p < .05$. LV: low valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.13

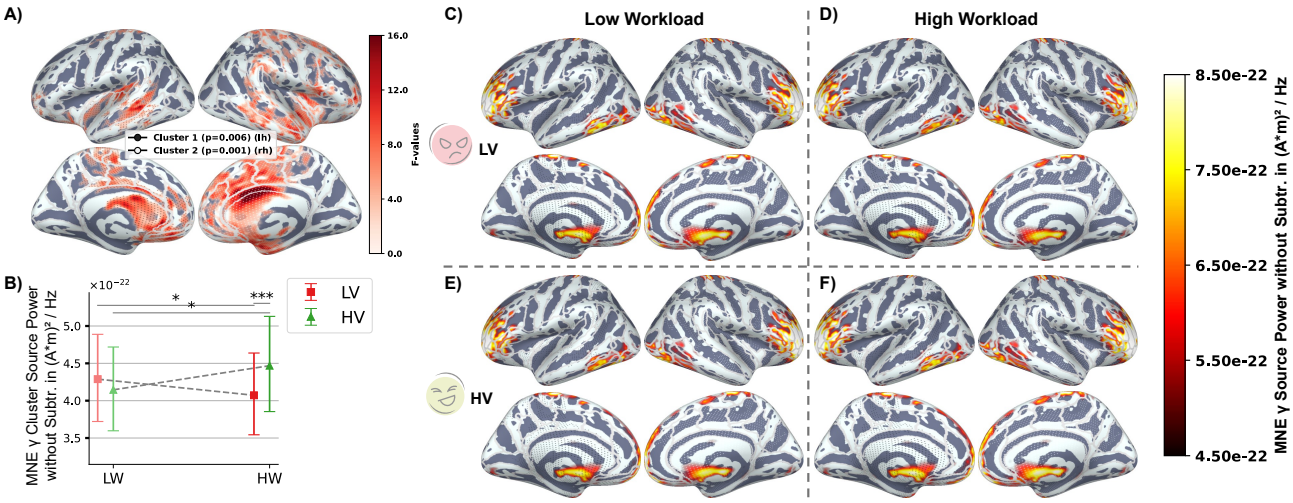
Permutation-Based Spatial Clustering Results of the Interaction Effect on the Aperiodic Broadband Component



Note. Panel (A) shows the spatial distribution of significant *F*-values of the clustering using the aperiodic broadband. Source power was estimated using the minimum norm estimation (MNE) method. Panel (B) shows the averaged aperiodic broadband per condition, averaged across significant vertices. Post-hoc comparisons were conducted using Wilcoxon signed-rank tests (FDR-corrected). Coloured symbols and error bars represent the bootstrapped grand averages for each condition, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI) across participants. Panels (C) to (F) display the aperiodic broadband component per condition in source space. Significance levels from the Wilcoxon signed-rank test: *** for $p < .001$, ** for $p < .01$, * for $p < .05$. LV: low valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary Figure 8.14

Permutation-Based Spatial Clustering Results of the Gamma-Band Interaction Effect Without Subtracting the Aperiodic Component



Note. Spatial distribution of significant *F*-values of the clustering is visualised in Panel (A). Clustering was performed on oscillatory gamma-band power without subtracting the aperiodic component. Source power was estimated using the minimum norm estimation (MNE) method. Panel (B) shows the gamma-band source power within significant vertices for the conditions. Post-hoc comparisons were conducted using Wilcoxon signed-rank tests (FDR-corrected). Coloured symbols and error bars represent the bootstrapped grand averages for each condition, along with their Bonferroni-corrected 2.5^{th} and 97.5^{th} percentile confidence intervals (CI) across participants. Panels (C) to (F) display the source gamma-band power per condition. Significance levels from the Wilcoxon signed-rank test: *** for $p < .001$, ** for $p < .01$, * for $p < .05$. LV: low valence; HV: high valence; LW: low visuo-spatial workload; HW: high visuo-spatial workload.

Supplementary References Study III

Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., & Müller, K.-R. (2008). Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Processing Magazine*, *25*(1), 41–56. https://doi.org/10.1109/MSP.2008.4408441

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, *23*(12), 1655–1665. https://doi.org/10.1038/s41593-020-00744-x

Duchowski, A. T. (2018). The index of pupillary activity: Measuring cognitive load vis-à-vis task difficulty with pupil oscillation. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, *282*, 1–13. https://doi.org/10.1145/3173574.3173856

Grosse-Wentrup, M., & Buss, M. (2008). Multiclass common spatial patterns and information theoretic feature extraction. *IEEE Transactions on Biomedical Engineering*, *55*(8), 1991–2000. https://doi.org/10.1109/TBME.2008.921154

Jacob, M. S., Roach, B. J., Sargent, K. S., Mathalon, D. H., & Ford, J. M. (2021). Aperiodic measures of neural excitability are associated with anticorrelated hemodynamic networks at rest: A combined EEG-fMRI study. *NeuroImage*, *245*, 118705. https://doi.org/10.1016/j.neuroimage.2021.118705

Lingelbach, K., Vukelić, M., & Rieger, J. W. (2024). GAUDIE: Development, validation, and exploration of a naturalistic German auditory emotional database. *Behavior Research Methods*, *56*, 2049–2063. https://doi.org/10.3758/s13428-023-02135-z

Lu, R., Dermody, N., Duncan, J., & Woolgar, A. (2024). Aperiodic and oscillatory systems underpinning human domain-general cognition. *Communications Biology*, *7*, 1643. https://doi.org/10.1038/s42003-024-07397-7

Thuwal, K., Banerjee, A., & Roy, D. (2021). Aperiodic and periodic components of ongoing oscillatory brain dynamics link distinct functional aspects of cognition across adult lifespan. *eNeuro*, *8*(5). https://doi.org/10.1523/ENEURO.0224-21.2021

Zhang, C., Stock, A.-K., Mückschel, M., Hommel, B., & Beste, C. (2023). Aperiodic neural activity reflects metacontrol. *Cerebral Cortex*, *33*(12), 7941–7951. https://doi.org/10.1093/cercor/bhad089

Publication List

I hereby confirm that Katharina Lingelbach contributed to the aforementioned studies as stated below:

Study I: Lingelbach, K., & Rieger, J. W. (2025). Neurophysiological basis of emotional face perception and working memory load in a dual-task MEG study. *Human Brain Mapping, 46*(8), 1–25. https://doi.org/10.1002/hbm.70242
Author contributions: Katharina Lingelbach: Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualisation, Writing – original draft, Writing – review & editing. Jochem W. Rieger: Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Study II: Lingelbach, K., Vukelić, M., & Rieger, J. W. (2024). GAUDIE: Development, validation, and exploration of a naturalistic German auditory emotional database. *Behavior Research Methods, 56*, 2049–2063. https://doi.org/10.3758/s13428-023-02135-z
Author contributions: Katharina Lingelbach: Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualisation, Writing – original draft, Writing – review & editing. Jochem W. Rieger: Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.

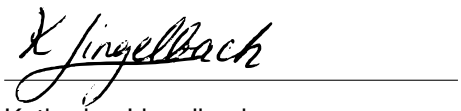
Study III: Lingelbach, K., Herrmann, C. S., & Rieger, J. W. (2026). Oscillatory and gaze signatures of socio-emotional speech processing, visuo-spatial cognition, and their interaction in a near-realistic dual-task MEG study. *Imaging Neuroscience, 4*, 1–27. https://doi.org/10.1162/IMAG.a.1134
Author contributions: Katharina Lingelbach: Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualisation, Writing – original draft, Writing – review & editing. Christoph S. Herrmann: Supervision, Validation, Writing – review & editing. Jochem W. Rieger: Conceptualisation, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.



Prof. Dr. Jochem W. Rieger
Oldenburg, 30.01.2026

Declaration of Academic Integrity

I completed this work independently and used only the indicated resources. This dissertation is entirely my own work. All the sources of information have been acknowledged by means of complete references. The dissertation has neither as a whole nor in part been published, except where explicitly stated in the publication list. The dissertation has neither as a whole nor in part been submitted for assessment in a doctoral procedure at another university. I am aware of the guidelines of good scientific practice of the Carl von Ossietzky University Oldenburg, and I have adhered to them in preparing this dissertation. I confirm that I have not availed myself of any commercial placement or consulting services in connection with my promotion procedure.



Katharina Lingelbach

Oldenburg, 30.01.2026

Declaration of Generative AI in Scientific Writing

In preparation of this thesis, I have used ChatGPT (https://chatgpt.com) and DeepL Write (https://www.deepl.com/de/write) to refine the text and improve readability. I declare that I have used the above-mentioned AI tools solely for spell-checking, grammar, and wording of this thesis. I have read the ”Recommendations for the Use of Generative AI in Doctoral Studies, Especially Dissertations, at the University of Oldenburg” (released on 12th November 2024) and have actively informed myself about the capabilities and limitations of the above-mentioned AI tools to the extent that I can use them responsibly. Following their use, I reviewed and edited the generated output and take full responsibility for the content of this thesis.



Katharina Lingelbach
Oldenburg, 30.01.2026

CURRICULUM VITAE

Katharina Lingelbach (née Hirning)

Kelterstraße 24, 70199 Stuttgart
Birthday: 29.10.1993, Germany
✉ Katharina.Lingelbach@mailbox.org
https://katharinalingelbach.com
https://orcid.org/0000-0003-1895-1667



PROFESSIONAL CAREER

starting 05.2026	**W1 Professorship in Neuroergonomics, University of Stuttgart**
02.2021 – today	**Fraunhofer Institute for Industrial Engineering IAO, Stuttgart** Research Associate in the Research Group Applied Neurocognitive Systems - Research in Neuroergonomics and Brain-Computer Interfaces (BCIs) - Acquisition of third-party funding - Supervision of student projects and theses - Specialised in M/EEG, fNIRS, and eye-tracking signal processing, multivariate statistics and machine learning for BCIs
11.2018 – 01.2021	**Institute of Human Factors and Technology Management IAT, University Stuttgart** Research Associate in collaboration with the Fraunhofer IAO Department Human-Technology Interaction

EDUCATION

02.2021 – 2025	**Department of Psychology, Carl von Ossietzky University Oldenburg, Germany** PhD Student in the Research Group Applied Neurocognitive Psychology (Dr. rer nat.) Doctoral Thesis *Neurophysiological basis of interacting emotional and cognitive processes* Supervision: Prof. Jochem Rieger, Prof. Christoph Herrmann - Conception, conduction and analysis of a MEG-eye-tracking study (*N* = 50) to investigate emotional face processing and working memory load in a dual-task Associated Publication: ▪ Lingelbach, K. & Rieger, J. W. (2025). Neurophysiological basis of emotional face perception and working memory load in a dual-task MEG study. *Human Brain Mapping, 46*(8), 1-25. - Development and validation of the naturalistic stimulus database GAUDIE Associated Publication: ▪ Lingelbach, K., Vukelić, M., & Rieger, J. W. (2024). GAUDIE: Development, validation, and exploration of a naturalistic German AUDItory Emotional database. *Behavior Research Methods, 56,* 2049-2063. - Conception, conduction and analysis of a MEG-eye-tracking study (*N* = 53) to investigate the effect of emotional speech during different task load levels in a driving simulation

Associated Publications:

- Lingelbach, K., Herrmann, C. S., & Rieger, J. W. (2026). Oscillatory and gaze signatures of socio-emotional speech processing, visuo-spatial cognition, and their interaction in a near-realistic dual-task MEG study. *Imaging Neuroscience, 4*, 1-27.

01.2018 – 10.2018 **Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany**

Master Student in the Research Group Adaptive Memory

10.2016 – 10.2018 **Department of Psychology, University of Vienna, Austria**

M.Sc. in Psychology – Mind & Brain (*Grade: 1,0)*

Master Thesis *The role of the right dorsolateral prefrontal cortex in voluntary retrieval suppression: An online rTMS approach (Thesis Grade: 1,0)*

Supervision: Prof. Ulrich Ansorge, Dr. Roland Benoit

Associated Publication:

- Stramaccia, D. F., Bergmann, F., Lingelbach, K., Numssen, O., Hartwigsen, G., & Benoit, R. G. (2025). Hindering memory suppression by perturbing the right dorsolateral prefrontal cortex. *Journal of Cognitive Neuroscience, 10*, 1-13.

10.2014 – 07.2016 **Department of Psychology, University of Vienna, Austria**

B.Sc. in Psychology *(Grade: 1,47)*

Bachelor Thesis *The pop-out effect for faces within visual search (Thesis Grade: 1,0)*
Supervision: Prof. Ulrich Ansorge

RESEARCH AWARDS AND SCHOLARSHIPS

06.2025 **Best Poster Award** Psychologie und Gehirn 2025, DGPs Fachgruppe Biologische Psychologie und Neuropsychologie

Contribution *Decoding the neural basis of interactions between emotional speech processing and attentional control in realistic environments*

09.2024 **Scholarship** Wolfgang Schulenberg Programm, Carl von Ossietzky University Oldenburg

07.2024 **Best Talk Award** 5th International Neuroergonomics Conference, Bordeaux, France

Contribution *Emotional detours: Oscillatory signatures of emotion and cognitive load during simulated driving*

07.2023 **Best Paper Award** AHFE 2023 16th International Conference, Section Neuroergonomics and Cognitive Engineering

Contribution *Investigating feature set decisions for mental state decoding in virtual reality-based learning environments*

10.2022 **Co-Author Best Talk Award** 3rd Neuroadaptive Technology Conference, Berlin, Germany

Contribution *BCI-based deep reinforcement learning for robot training*

09.2021 **Best Talk Award** 3rd International Neuroergonomics Conference, Munich, Germany

Contribution *Investigating the emotion-cognition interaction: Effects of affective distractors on working memory load*

02.2021 – 02.2023 **Research program** Fraunhofer TALENTA start

10.2014 – 10.2018 **Scholarship** for Outstanding Performance University of Vienna

INVITED TALKS AND ORGANIZED SYMPOSIA

05.2024 **Invited Talk**, Research Colloquium, Institute of Psychology, Julius-Maximilians-Universität Würzburg, Germany

Shared resources and interacting brain networks of emotional and cognitive processes

04.2024 **Keynote** at the NIRx fNIRS Workshop 2024, Berlin, Germany

Brain-computer interfaces in everyday life: Mental state encoding and decoding with functional near-infrared spectroscopy

10.2024 **Co-organiser of the Symposium** *M/EEG correlates of emotional relevance: Perceptual versus psychological contextual effects* at the Annual Meeting of the Society for Psychophysiological Research, Prague, Czech Republic (Chair by Sebastian Schindler)

07.2024 **Organiser and Session Chair of the Symposium** *Towards the study of interacting emotional and cognitive processes and their neurophysiological basis* at the Conference Psychologie & Gehirn, Hamburg, Germany

11.2019 **Co-Organizer** of the Event »*Interaktion mit Hirn*«, Fraunhofer IAO, Stuttgart, Germany

THIRD-PARTY FUNDING [TOTAL: 326.600 €]

2020 – 2022 **MINT-Innovationen, Vector Stiftung**

Funding amount: 100.000 €

RomEO – Robotik mit emotions-basiertem reinforcement learning

Role: Grant application and principal investigator

Associated publication:

- Lingelbach, K., Rips, J., Karstensen, L., Mathis-Ullrich, F., & Vukelić, M. (2025). Evaluating robotic actions: Spatiotemporal brain dynamics of performance assessment in robot-assisted laparoscopic training. *Frontiers in Neuroergonomics, 6,* 1535799.

2020 – 2022 **ZIM-Project, German Federal Ministry for Economic Affairs and Energy**

Funding amount: 190.000 €

NIRcademy – Mobile Neuro-Bildgebung für virtuelle Lernumgebungen

Role: Project lead and contributed to the grant application

Associated publications:

- Lingelbach, K., Diers, D., & Vukelić, M. (2023). Towards user-aware VR learning environments: Combining brain-computer interfaces with virtual

reality for mental state decoding. In *2023 CHI conference on human factors in computing systems,* 292, Association for Computing Machinery.

- Lingelbach, K., Diers, D., Bui, M., & Vukelić, M. (2023). Investigating feature set decisions for mental state decoding in virtual reality-based learning environments. In *Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2023 International Conference on Neuroergonomics and Cognitive Engineering, July 20–24, 2023, San Francisco, USA*.

2022 – 2022 **Subcontract within a collaborative project at the University of Stuttgart**

Amount: 36.600

MIkado – Multimodale Indikatorisierung kognitiver und affektiver Prozesse in domänenübergreifenden Lernverläufen

Role: Cooperation partner of LLiS, University of Stuttgart, and project collaborator

Associated publications:

- Lingelbach, K., Gado, S., Wirzberger, M., & Vukelić, M. (2023). Workload-dependent hemispheric asymmetries during the emotion-cognition interaction: A close-to-naturalistic fNIRS study. *Frontiers in Neuroergonomics, 4,* 1273810.
- Gado, S.*, Lingelbach, K.*, Wirzberger, M., & Vukelić, M. (2023). Decoding mental effort in a quasi-realistic scenario: A feasibility study on multimodal data fusion and classification. *Sensors, 23*, 6546.

PATENTS AND INVENTION DISCLOSURE

22.09.2022 **Patent ID:** DE102021106689 (A1)

Verfahren und System zur Vigilanzerhöhung und/oder kognitiven Leistungssteigerung eines Fahrzeuginsassen und Fahrzeug mit dergleichen

Brief description: The invention describes a system utilising non-perceptible rhythmic light stimulation to enhance vigilance and cognitive performance of vehicle occupants.

06.05.2021 **Patent ID:** WO2021083512A1

Measuring an attentional state and providing automatic feedback during a technical system interaction

Brief description: The invention relates to an approach to improving a user's attention and focus by measuring the current attentional state while interacting with a technical system and providing automatic feedback.

TEACHING AND SUPERVISION

2021 – 2024 **Seminar and supervision**

Course *Practical Projects* (9 ECTS, Summer semester), Master's programme Neurocognitive Psychology, Department of Psychology, Carl von Ossietzky University Oldenburg

2020 – Sebastian Sander: *Multivariate pattern analysis: An age comparison of temporal dynamics underlying emotional states*

2021 – Sumbul Jafri: *Assessing effects of positive emotional speech distractors on increasing mental workload with fnirs*

2021 – Alina Ruge: *Personality traits and emotional face processing at different working memory load levels*

2021 – Nikolina Polimac: *The role of dominance, familiarity, and personality traits in valence and arousal ratings of emotional speech*

2021 – Melanie Klapprott: *Neuroergonomic indices of emotional and cognitive processes during simulated driving*

2024 – Hatice Ünlü: *How do task difficulty and emotional speech affect gaze behaviour during simulated driving?*

SS 2022 **Course teaching**

Lecture substitute *Applied Cognitive Psychology* (6 ECTS), Master's programme Neurocognitive Psychology, Department of Psychology, Carl von Ossietzky University Oldenburg

Since 2019 **Thesis supervision in cooperation with the Institute of Human Factors and Technology Management (IAT), University of Stuttgart**

2024 - 2025 Master's thesis Pema Sherpa: *Assessing cognitive load in a driving simulator using blink-evoked responses with magnetoencephalography*. Neuroscience, Department of Neuroscience, Carl von Ossietzky University Oldenburg

2022 - 2023 Master's thesis Jennifer Rips: *Untersuchung von ereigniskorrelierten Potentialen während der Leistungsüberwachung von Roboteraktionen.* Psychology, Institute for Experimental Psychology, Heinrich Heine University Düsseldorf

2021 - 2022 Bachelor's thesis Sabrina Gado: *Decoding cognitive load using neurophysiological data.* Business Informatics, Faculty of Computer Science and Business Informatics, University of Applied Sciences Würzburg-Schweinfurt

2019 - 2020 Master's thesis Sebastian Sander: *Comparison of feature spaces for the prediction of working memory levels from fNIRS in a visuo-spatial virtual reality n-back paradigm.* Neurocognitive Psychology, Department of Psychology, Carl von Ossietzky University Oldenburg

MEMBERSHIPS AND ACTIVITIES IN THE RESEARCH COMMUNITY

Since 2024 Member of the Society for Psychophysiological Research SPR

Since 2024 Contributing Developer for mne-python and mne-nirs

Since 2023 Member of the CuttingEEG Association: A collective for Cutting-edge EEG/MEG research

Since 2022 Member of the Deutsche Gesellschaft für Psychologie (DGP)

Section Biological Psychology and Neuropsychology

07.2020 Certified Data Scientist in Machine Learning and Deep Learning

Fraunhofer BIG DATA AI *(Exam Grade 97 %)*

Ad-hoc Reviewer (Selection) Human Brain Mapping, Scientific Reports, European Journal of Neuroscience, Frontiers of Human Neuroscience, Frontiers of Psychology, Frontiers of Neuroergonomics, Applied Ergonomics

PUBLICATIONS

Peer-Reviewed Journal Articles

Lingelbach, K., Herrmann, C. S., & Rieger, J. W. (2026). Oscillatory and gaze signatures of socio-emotional speech processing, visuo-spatial cognition, and their interaction in a near-realistic dual-task MEG study. *Imaging Neuroscience, 4*, 1-27. https://doi.org/10.1162/IMAG.a.1134

Vorreuther, A., Tagalidou, N., Lingelbach, K., Hubert, A., Bareiß, A., Nittel, T., …, Vukelić, M. (2026). Response inhibition in autistic adults: A functional near-infrared spectroscopy study in virtual reality. *Brain and Behavior* [accepted, in print]. https://doi.org/10.1002/brb3.71249

Lingelbach, K., Vorreuther, A., Moll, E., & Vukelić, M. (2025). Shielding the mind with flow: Attention allocation and auditory event-related potentials under varying mental workload. *European Journal of Neuroscience, 62(8).* https://doi.org/10.1111/ejn.70283

Stramaccia, D. F., Bergmann, F., Lingelbach, K., Numssen, O., Hartwigsen, G., & Benoit, R. G. (2025). Hindering memory suppression by perturbing the right dorsolateral PFC. *Journal of Cognitive Neuroscience*, 1–13. https://doi.org/10.1162/jocn.a.90

Lingelbach, K., & Rieger, J. W. (2025). Neurophysiological basis of emotional face perception and working memory load in a dual-task MEG study. *Human Brain Mapping, 46*(8), 1–25. https://doi.org/10.1002/hbm.70242

Lingelbach, K., Rips, J., Karstensen, L., Mathis-Ullrich, F., & Vukelić, M. (2025). Evaluating robotic actions: Spatiotemporal brain dynamics of performance assessment in robot-assisted laparoscopic training. *Frontiers in Neuroergonomics, 6*, 1535799. https://doi.org/10.3389/fnrgo.2025.1535799

Lingelbach, K., Vukelić, M., & Rieger, J. W. (2024). GAUDIE: Development, validation, and exploration of a naturalistic German AUDItory Emotional database. *Behavior Research Methods, 56*, 2049–2063. https://doi.org/10.3758/s13428-023-02135-z

Caballero-Bruno, I., Lingelbach, K., Wohllebe, T., Weng, M., Piechnik, D., Tagalidou, N., ... & Hernández-Castellano, P. M. (2024). Sleep quality and comfort in fully automated vehicles: A comparison of two seat configurations. *Applied Ergonomics, 114*, 104137. https://doi.org/10.1016/j.apergo.2024.104137

Lingelbach, K., Gado, S., Wirzberger, M., & Vukelić, M. (2023). Workload-dependent hemispheric asymmetries during the emotion-cognition interaction: A close-to-naturalistic fNIRS study. *Frontiers in Neuroergonomics, 4*, 1273810. https://doi.org/10.3389/fnrgo.2023.1273810

Gado, S.*, Lingelbach, K.*, Wirzberger, M., & Vukelić, M. (2023). Decoding mental effort in a quasi-realistic scenario: A feasibility study on multimodal data fusion and classification. *Sensors, 23*(14), 6546. https://doi.org/10.3390/s23146546

Vukelić, M., Bui, M., Vorreuther, A., & Lingelbach, K. (2023). Combining brain-computer interfaces with deep reinforcement learning for robot training: A feasibility study in a simulation environment. *Frontiers in Neuroergonomics, 4*, 1274730. https://doi.org/10.3389/fnrgo.2023.1274730

Lingelbach, K., Dreyer, A. M., Schöllhorn, I., Bui, M., Weng, M., Diederichs, F., ... & Vukelić, M. (2021). Brain oscillation entrainment by perceptible and non-perceptible rhythmic light stimulation. *Frontiers in Neuroergonomics, 2*, 646225. https://doi.org/10.3389/fnrgo.2021.646225

Gado, S., Kempen, R., Lingelbach, K., & Bipp, T. (2021). Artificial intelligence in psychology: How can we enable psychology students to accept and use artificial intelligence? *Psychology Learning & Teaching, 21*(1), 37–56. https://doi.org/10.1177/14757257211037149

Lingelbach, K., Piechnik, D., Gado, S., Janssen, D., Eichler, M., Hentschel, L., ... & Peissner, M. (2021). Effects of the COVID-19 pandemic on psychological well-being and mental health based on a German online survey. *Frontiers in Public Health, 9*, 655083. https://doi.org/10.3389/fpubh.2021.655083

Vukelić, M., Lingelbach, K., Pollmann, K., & Peissner, M. (2020). Oscillatory EEG signatures of affective processes during interaction with adaptive computer systems. *Brain Sciences, 11*(1), 35. https://doi.org/10.3390/brainsci11010035

Conference Proceedings

Lingelbach, K., Diers, D., Bui, M., & Vukelić, M. (2023). Investigating feature set decisions for mental state decoding in virtual reality-based learning environments. In *Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2023 International Conference on Neuroergonomics and Cognitive Engineering*, July 20–24, 2023, San Francisco, USA. AHFE Open Access. https://doi.org/10.54941/ahfe1003014

Lingelbach, K., Diers, D., & Vukelić, M. (2023). Towards user-aware VR learning environments: Combining brain-computer interfaces with virtual reality for mental state decoding. In *2023 CHI conference on human factors in computing systems* (Article 292, 8 pages). Association for Computing Machinery. https://doi.org/10.1145/3544549.3585716

Lingelbach, K., Tagalidou, N., Markey, P., Föll, B., Peissner, M. & Vukelic, M. (2022). Examining joy of use and usability during mobile phone interactions within a multimodal methods approach. In *Proceedings of Mensch Und Computer 2022* (pp. 276–285). Association for Computing Machinery. https://doi.org/10.1145/3543758.3543771

Lingelbach, K., Gado, S., Janssen, D., Piechnik, D., Eichler, M., Knopf, D., Hentschel, L., Schuler, M., Sernatinger, D., & Peissner, M. (2022). Identifying the Effects of COVID-19 psychological well-being through unsupervised clustering for mixed data. In X.-S. Yang, S. Sherratt, N. Dey, & A. Joshi (Eds.), *Proceedings of Sixth International Congress on Information and Communication Technology* (pp. 883–895). Springer, Singapore. https://doi.org/10.1007/978-981-16-2377-6_81

Brouwer, A. M., Stuldreher, I., Penen, S. H., Lingelbach, K., & Vukelić, M. (2021, October 15–18). Combining eye tracking and physiology for detection of emotion and workload. In A. Spink, J. Barski, A.-M. Brouwer, G. Riedel, & A. Sil (Eds.), *Proceedings of the joint 12th International Conference on Methods and Techniques in Behavioral Research and the 6th Seminar on Behavioral Methods* (Vol. 1, pp. 1–12). https://doi.org/10.6084/m9.figshare.13013717

Lingelbach, K., Schöllhorn, I., Dreyer, A. M., Diederichs, F., Bui, M., Weng, M., Rieger, J. W., Petermann-Stock, I., & Vukelić, M. (2021). Investigating the modulation of spatio-temporal and oscillatory power dynamics by perceptible and non-perceptible rhythmic light stimulation. In H. Ayaz, U. Asgher, & L. Paletta (Eds.), *Advances in Neuroergonomics and cognitive engineering: Proceedings of the AHFE 2021 virtual conferences on Neuroergonomics and cognitive engineering, industrial cognitive ergonomics and engineering psychology, and cognitive computing and internet of things, July 25–29, 2021, USA* (pp. 11–19). Springer. https://doi.org/10.1007/978-3-030-80285-1_2

Lingelbach, K., Gado, S., & Bauer, W. (2021). Neuro-adaptive tutoring systems: Neurophysiological-based recognition of affective-emotional and cognitive states of learners for intelligent neuro-adaptive tutoring systems. In *Schriftenreihe der Wissenschaftlichen Gesellschaft für Arbeits- und Betriebsorganisation (WGAB) e.V.(2021),* (p. 243–260). https://doi.org/10.30844/wgab_2021_15

Gado, S.*, Lingelbach, K.*, Bui, M., Rieger, J. W., & Vukelic, M. (2021). Real-time feedback of subjective affect and working memory load based on neurophysiological activity. In *HCI International 2021-Late Breaking Posters: 23rd HCI International Conference*, *HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part II* (pp. 80–87). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-90179-0_11

Lingelbach, K., Bui, M., Diederichs, F., & Vukelić, M. (2021). Exploring conventional, automated and deep machine learning for electrodermal activity-based drivers' stress recognition. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 1339–1344). IEEE. https://doi.org/10.1109/SMC52423.2021.9658875

Lingelbach, K., Lingelbach, Y., Otte, S., Bui, M., Künzell, T., & Peissner, M. (2021). Demand forecasting using ensemble learning for effective scheduling of logistic orders. In T. Z. Ahram, W. Karwowski, & J. Kalra (Eds.), *Advances in artificial intelligence, software and systems engineering* (Vol. 271). Springer. https://doi.org/10.1007/978-3-030-80624-8_39

Kosuru, R. K.*, Lingelbach, K.*, Bui, M., & Vukelić, M. (2019). Mindtrain: How to train your mind with interactive technologies. In *Proceedings of Mensch und Computer 2019* (pp. 643–647). Association for Computing Machinery. https://doi.org/10.1145/3340764.3344884

Books, Book Chapters and Monographs

Amthor, F. (2024). *Das menschliche Gehirn für Dummies* (3., vollständig überarb. Aufl. von K. Lingelbach). Wiley-VCH. ISBN 978-3-527-72176-4

Vukelić, M., Lingelbach, K., & Piechnik, D. (2021). Feinfühlige Technik: Wie Neuroergonomie und Brain-Computer-Interfaces in der Praxis eingesetzt werden können. In W. Bauer, O. Riedel, T. Renner, & M. Peissner (Eds.), Stuttgart: Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO. https://doi.org/10.24406/publica-fhg-300856

Lis, D., Tagalidou, N., Lingelbach, K., & Spiekermann, M. (2019). *Ökosysteme für Daten und Künstliche Intelligenz: Positionspapier*. Fraunhofer-Gesellschaft. ISBN 978-3-00-064163-3. https://doi.org/10.24406/ISST-N-543753

Talks and Conference Contributions

Lingelbach, K., Herrmann, C. S. & Rieger, J. W. (2025, June 18–21). Neural and pupillary oscillations of external and internal information processing modes during emotional speech processing and visuo-spatial workload [Poster contribution]. Psychologie und Gehirn 2025, Biological Psychology and Neuropsychology, German Psychological Society (DGP), Julius-Maximilians-University of Würzburg, Germany.

Lingelbach, K., & Rieger, J. W. (2025, June 18–21). Decoding the neural basis of interactions between emotional speech processing and attentional control in realistic environments [Poster contribution]. Psychologie und Gehirn 2025, Biological Psychology and Neuropsychology, German Psychological Society (DGP), Julius-Maximilians-University of Würzburg, Germany.

Vukelić, M., Bui, M., Lingelbach, K., & Tagalidou, N. (2025, April 13–15). *Balancing the gamer's load: Blink-driven EEG for neuroadaptive frameworks* [Symposium contribution]. Methods in Mobile EEG Conference, Belgrade, Serbia.

Lingelbach, K., & Rieger, J. W. (2024, October 23–26). *Spatiotemporal and Oscillatory Signatures of Emotional Face Processing and Working Memory Load in a Dual-Task* [Symposium contribution]. 64th Annual Meeting of the Society for Psychophysiological Research, Prague, Czech Republic.

Lingelbach, K., Herrmann, C. S., & Rieger, J. W. (2024, October 23–26). Navigating through emotion-cognition interactions: Brain oscillations of visuo-spatial attention and auditory emotional processing during simulated driving [Poster contribution]. 64th Annual Meeting of the Society for Psychophysiological Research, Prague, Czech Republic.

Lingelbach, K., & Rieger, J. W. (2024, September 12–15). *Decoding the interaction between emotional and cognitive processes with functional near-infrared spectroscopy*. Paper presented at the fNIRS 2024 Conference – Society for Functional Near-Infrared Spectroscopy, Birmingham, United Kingdom.

Lingelbach, K., & Rieger, J. W. (2024, July 8–12). *Emotional detours: Oscillatory signatures of emotion and cognitive load during simulated driving* [Symposium contribution]. 5th International Neuroergonomics Conference, Bordeaux, France.

Lingelbach, K., Rips, J., & Vukelić, M. (2024, July 8–12). *Bridging brains and bots: Spatiotemporal signatures of evaluative robot observations* [Poster contribution]. 5th International Neuroergonomics Conference, Bordeaux, France.

Lingelbach, K. (2024, June 2–5). *Our mind at work: fNIRS-based encoding and decoding of cognitive load and learning dynamics in simulated environments* [Symposium contribution, presented by Vukelić, M.]. 5th International Mobile Brain/Body Imaging (MoBI) Conference, Piran, Slovenia.

Lingelbach, K. (2024, May 29–June 1). *Towards the study of interacting emotional and cognitive processes and their neurophysiological basis* [Symposium chair abstract]. Psychologie und Gehirn 2024, Biological Psychology and Neuropsychology, German Psychological Society (DGP), University of Hamburg, Germany.

Lingelbach, K. & Rieger, J.W. (2024, May 29–June 1). *Spatiotemporal and oscillatory signatures of emotional face processing and working memory load in a dual-task* [Symposium contribution]. Psychologie und Gehirn 2024, Biological Psychology and Neuropsychology, German Psychological Society (DGP), University of Hamburg, Germany.

Vorreuther, A., & Lingelbach, K. (2024, May 29–June 1). *fNIRS-based decoding of mental state in virtual reality* [Symposium contribution]. Psychologie und Gehirn 2024, Biological Psychology and Neuropsychology, German Psychological Society (DGP), University of Hamburg, Germany.

Lingelbach, K. (2023, June 8–10). *Investigating learning processes with functional near-infrared spectroscopy and virtual reality* [Symposium contribution]. Psychologie und Gehirn 2023, Biological Psychology and Neuropsychology, German Psychological Society (DGP), University of Tübingen, Germany.

Lingelbach, K. (2023, June 8–10). *GAUDIE: Development, validation and exploration of a naturalistic German AUDItory Emotional Database* [Poster contribution]. Psychologie und Gehirn 2023, Biological Psychology and Neuropsychology, German Psychological Society (DGP), University of Tübingen, Germany.

Lingelbach, K., Gado, S., Vukelić, M., & Rieger, J. W. (2021, September 11–16). *Investigating the emotion-cognition interaction: Effects of affective distractors on working memory load.* [Symposium contribution]. 3rd International Neuroergonomics Conference on The Brain at Work and in Everyday Life, Munich, Germany. https://neuroergonomicsconference.um.ifi.lmu.de/wp-content/uploads/submissions/120.pdf

Lingelbach, K., Gado, S., Vukelić, M., & Rieger, J. W. (2021, September 11–16). *What I feel and what I say: Decoding neurophysiological correlates of cognitive and affective states* [Poster contribution]. 3rd International Neuroergonomics Conference on The Brain at Work and in everyday life, Munich, Germany. https://neuroergonomicsconference.um.ifi.lmu.de/wp-content/uploads/submissions/126.pdf

Lingelbach, K., Piechnik, D., Brouwer, A.-M., Stuldreher, I., & Vukelić, M. (2021, September 11–16). *Temporal decoding of emotion and workload from fixation-related EEG recordings* [Poster contribution]. 3rd International Neuroergonomics Conference on The Brain at Work and in everyday life, Munich, Germany. https://neuroergonomicsconference.um.ifi.lmu.de/wp-content/uploads/submissions/107.pdf

Lingelbach, K. (2020, November 12). *Combining neuroscience and machine learning for a human-centered AI approach: Decoding of cognitive and affective states* [Conference presentation]. World Usability Day (WUD) 2020 Stuttgart, Online, Germany.

Lingelbach, K., Peissner, M., & Vukelić, M. (2020, September 21–22). *Towards biointelligent closed-loop systems: Hybrid brain-computer interfaces for the decoding of cognitive and affective states during human-machine interaction* [Conference presentation]. Bioeconomy Innovation Congress 2020, University of Hohenheim, Online, Germany.